

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 19:52:25 ; Search time 4776.29 Seconds

(without alignments)
11254.112 Million cell updates/sec

Title: US-08-876-132-2

Perfect score: 1847

Sequence: 1 ACATCTCAACACAGTTTAAAA.....AAACAGCAAAAAAGACATCT 1847

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Genemdb1:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pal:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1066.6	57.7	3661	1	AF128889
2	130	7.0	7597	6	AX345915
3	123.2	6.7	6286	6	AX251345
4	114.2	6.2	6109	6	AX251772
5	114.2	6.2	6109	6	AX345228
6	114.2	6.2	11861	9	AC069435
7	113.8	6.2	7814	6	AX251284
8	112.6	6.1	171105	3	AC007532
9	112.6	6.1	190801	3	AC009212
10	111.6	6.0	12237	6	AX347260
11	111	6.0	294914	3	AE003603
12	110	6.0	67970	3	PFMAL1P3
13	109	5.9	14867	3	AE001398
14	108.8	5.9	90487	9	AL592166
15	106.8	5.8	213056	9	AL103590
16	105.2	5.7	161078	9	AP003500
17	105.2	5.7	161286	9	AC025120
18	103.4	5.6	100925	9	AC012627
19	103.4	5.6	120481	9	AC099738
20	103.4	5.6	177293	2	AC069525
21	103.4	5.6	235532	9	AC008739
22	103.2	5.6	79018	9	AL627107
23	102.6	5.6	151032	9	AC092362
24	102.4	5.5	262608	2	AC092362
25	101.8	5.5	131682	9	AL672277
26	101.8	5.5	189082	9	AL7231537
27	101.2	5.5	349980	6	AX344535
28	100.8	5.5	170141	2	AL611928
29	100.8	5.5	201312	2	AC067900
30	99.8	5.4	50612	9	AL513530
31	99.8	5.4	81696	9	AF438327-3
32	99.8	5.4	86827	3	PFMAL3P5
33	99.8	5.4	110000	9	AF438327-2
34	99.6	5.4	17443	3	AC024806
35	99.6	5.4	150119	2	AL845366
36	99	5.4	155702	9	AC105451
37	98.4	5.3	136098	9	AC006970
38	98.4	5.3	170141	2	AL611928
39	98.2	5.3	148525	2	AC016675
40	97.8	5.3	13321	6	AX251175
41	97.8	5.3	158950	9	AL391425
42	97.6	5.3	86975	2	AC048340
43	97.4	5.3	162445	9	AL158151
44	97.2	5.3	124057	9	AP000770
45	97	5.3	5413	6	AX251450

ALIGNMENTS

RESULT 1
AF128889
LOCUS
DEFINITION Pantoea citrea strain ATCC31623 plasmid pPZG500, complete plasmid
ACCESSION AF128889
VERSION AF128889.1 GI:4457229
KEYWORDS
SOURCE Pantoea citrea.
ORGANISM Pantoea citrea.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
- Pantoea.
REFERENCE 1 (bases 1 to 3661)
AUTHORS Billic.M. and Delic.V.
TITLE Isolation and characterization of a cryptic plasmid from Erwinia

Pred. No. is the number of results predicted by chance to have a

JOURNAL J. Appl. Microbiol. 83 (4), 485-492 (1997)
MEDLINE 98012514
PUBMED 9351229
REFERENCE 2 (bases 1 to 3661)
AUTHORS Bilic Nezic, M. and Delic, V.
TITLE Sequence Analysis and Functional Characterization of Replication Region in Plasmid pPZG500 from *Pantoea citrea* ATCC 31623
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3661)
AUTHORS Bilic Nezic, M. and Delic, V.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-1999) Biotechnology, PUIVA d.d., Research Institute, Pilaz baruna Filipovica 25, Zagreb 10 000, Croatia
FEATURES
source
1. 3661
/organism="Pantoea citrea"
/strain="ATCC31623"
/db_xref="ATCC:31623"
/db_xref="taxon:5336"
/plasmid="pPZG500"
complement(join(3188..3661,1..177))
/note="RNAI"
99..317
/note="Orf1"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAD21205.1"
/db_xref="GI:4457230"
/translation="MGCAQCLCGSNHPRYSIGALAKLKHKLIOGLGFLYOKKKIK
QRKIFERNNILEFFRGDTCYRFAIFKI"
177
/note="Orf1V"
complement(334..342)
/note="putative"
/bound_moiety="Dnaa"
458..472
/bound_moiety="IHF"
489..704
/note="Orf2"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAD21206.1"
/db_xref="GI:4457231"
/translation="MDCVKKALENEIHFYKKGDIKSFLOLHYFDIDKALNGDEG
DIINSLSIDSLIFLMISOSAGLSIK"
complement(845..1066)
/note="Orf3"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAD21207.1"
/db_xref="GI:4457232"
/translation="MPSGSGSIGSYDIFNFISCKSSSINFSSTLKHRSQIIRAD
FLYLLEFRNOYLILFYISYGFISCVII"
complement(1193..1348)
/note="Orf4"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAD21208.1"
/db_xref="GI:4457233"
/translation="MODLFCWHLFPLHPIFYTHHFLFLYLKFSMMKFLSTERAENI
SHIFLKK"
1365
/note="Orf5"
/codon_start=1
/transl_table=1
/product="putative heat-shock protein"
/protein_id="AAD21209.1"
/db_xref="GI:4457234"

translation="MDEKSRKLTLENEKDLKIEASELKKKLTGPOGVLENTMKE
MKMNINDVNEETAGQYRKLFKNKVEHSKSDLYTGLCECTRNSFDKTRSAFRCIC
ERIQINDIEFTTKILQYKEKO"
2362..2991
/function="putative resolution of catenane plasmid dimers
prior to plasmid segregation to the daughter cells"
/note="Orf6"
/codon_start=1
/transl_table=1
/product="lambda-integrase like protein"
/protein_id="AAD21210.1"
/db_xref="GI:4457235"
/translation="WKEADYIMODIFKRLKNNKSYDRYAGFLSTCSTIGCPAEYLGK
IEIVRNREDEGISFKLILGAKVNDGQSEKRLHEDLSKHNDENMNTLSQDKNKEF
YKPDGKLYNSLROYLYIQHRTFSLTLHNRASDLKASGADDFIYIAFLGHRVTSOE
FYGARSSKGGIATVGVESDVVANKSQFAVSRTPSQISTLSKHFKE"
3154..3159
-35.signal 3175..3180
-10.signal 3188..3288
misc-RNA
/note="RNAI"
-10.signal 3295..3300
-35.signal 3311..3316
BASE COUNT 1366 a 472 c 730 g 1093 t
ORIGIN
Query Match 57.7%; Score 1066.6; DB 1; Length 3661;
Best Local Similarity 98.9%; Pred. No. 6.3e-157;
Matches 1128; Conservative 0; Mismatches 4; Indels 9; Gaps 5;
QY 710 AAGCTTGCGCTTCGTGGGTGCAAGCGTCAAGTGCATGCTGTGTCTTACTGTGTA 769
DB 1 AAGCTTGCGCTTCGTGGGTGCAAGCGTCAAGTGCATGCTGTGTGTCTTACTGTGTA 60
QY 770 ACTTTCACCTTTTTCGAGTGAATTCAGAGGCGCAACATGAGTCAAGCCAGCTTTCCT 829
DB 61 ACTTTCACCTTTTTCGAGTGAATTCAGAGGCGCAACATGAGTCAAGCCAGCTTTCCT 120
QY 830 GGGGTTCGCGACATCCAGCTTACAGCATTTGCTGCTTCGCAAGCTGAAGCAAAATTC 889
DB 121 GGGGTTCGCGACATCCAGCTTACAGCATTTGCTGCTTCGCAAGCTGAAGCAAAATTC 180
QY 890 TAAATCCAGGTTTGGGCTTTTATACGAGAGCAAAACAAAATAAACAAGAAAAA 949
DB 181 TAAATCCAGGTTTGGGCTTTTATACGAGAGCAAAACAAAATAAACAAGAAAAA 240
QY 950 TTTTCGACGCAAAAAATATTTTGAATTTTAAAGCGATCTTGCTACCGCACTTTTG 1009
DB 241 TTTTCGACGCAAAAAATATTTTGAATTTTAAAGCGATCTTGCTACCGCACTTTTG 300
QY 1010 CCAATATTTAAACCTGACTATCTTATATAGTTAA--TAGATATATCCGTTAGATATAA 1066
DB 301 CCAATATTTAAACCTGACTATCTTATATAGTTAA--TAGATATATCCGTTAGATATAA 360
QY 1067 AGTATGTTAAAGAGATTAACCAATTAATATATATATATATATATATATATATATAT 1126
DB 361 AGTATGTTAAAGAGATTAACCAATTAATATATATATATATATATATATATATATAT 420
QY 1127 AGTATGTTAAAGAGATTAACCAATTAATATATATATATATATATATATATATATAT 1186
DB 421 AGTATGTTAAAGAGATTAACCAATTAATATATATATATATATATATATATATATAT 480
QY 1187 GTAAATATATGATTTGATTTTAAAGAGATTAAGAAATGAATAGAACATTAATATAA 1246
DB 481 GTAAATATATGATTTGATTTTAAAGAGATTAAGAAATGAATAGAACATTAATATAA 540
QY 1247 AAGACGGGATATCAATCTTCTTACAATCTTGCAATCTTACTTGTATAGATAAGCAT 1306
DB 541 AAGACGGGATATCAATCTTCTTACAATCTTGCAATCTTACTTGTATAGATAAGCAT 600
QY 1307 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
DB 601 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659

QY 1367 AACTCTTGATGTGACCAATTCGGCTGGCTTTCATATAAATATACAGACGAA 1426
|||||
Db 660 AACTCTTGATGTGACCAATTCGGCTGGCTTTCATATAAATATACAGACGAA 716
QY 1427 ATGAATATTTATCATCAGCTAAACCTGAAATGATTTTAAAAAATACCTGCTTTATTC 1486
|||||
Db 717 ATGAATATTTATCATCAGCTAAACCTGAAATGATTTTAAAAAATACCTGCTTTATTC 776
QY 1487 ATTCGATCAATGGGAAAGATTTAATTAGCATGAGTGCATACATCAAGAAATGA 1546
|||||
Db 777 ATTCGATCAATGGGAAAGATTTAATTAGCATGAGTGCATACATCAAGAAATGA 836
QY 1547 TTAAGGATTTAGATATTTACATATGGAATAACCATATGATATTAATTAATTAAGA 1606
|||||
Db 837 TTAAGGATTTAGATATTTACATATGGAATAACCATATGATATTAATTAATTAAGA 896
QY 1607 ATATATGATTTTGAATATACAAAGATTTAAAGAAAGTCACTCTTAAAAATATCTGC 1666
|||||
Db 897 ATATATGATTTTGAATATACAAAGATTTAAAGAAAGTCACTCTTAAAAATATCTGC 956
QY 1667 ATTCGATCAATGGGAAAGATTTAATTAGCATGAGTGCATACATCAAGAAATGA 1726
|||||
Db 957 ATTCGATCAATGGGAAAGATTTAATTAGCATGAGTGCATACATCAAGAAATGA 1015
QY 1727 GATATTTGAAATATGCTATGAAATCTCTATGATGATCCAGATGATGATTAAGAAATC 1786
|||||
Db 1016 GATATTTGAAATATGCTATGAAATCTCTATGATGATCCAGATGATGATTAAGAAATC 1075
QY 1787 ACATATTCGACAGAGAAATGCTATGCTTTCCGATTTAAAAACAAGCAAAAAGATC 1846
|||||
Db 1076 ACATATTCGACAGAGAAATGCTATGCTTTCCGATTTAAAAACAAGCAAAAAGATC 1134
QY 1847 T 1847
Db 1135 T 1135

RESULT 2
AX345915/c
LOCUS AX345915 7597 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 986 from Patent WO0200928.
ACCESSION AX345915
VERSION AX345915.1 GI:18493801

KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.

REFERENCE
1
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 986 03-JAN-2002;
EpiGenomics AG (DE)

FEATURES
source
1..7597
Location/Qualifiers

/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 1965 a 109 c 1790 g 3733 t
ORIGIN

Query Match 7.0%; Score 130; DB 6; Length 7597;
Best Local Similarity 47.7%; Pred. No. 3.7e-11;
Matches 447; Conservative 0; Mismatches 485; Indels 6; Gaps 2;

QY 912 ATACCAAGAAAGCAAAATTAATTAACAAAGAAATTTTCAGCGAATAATATTTT 971
|||||
Db 1062 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1003
QY 972 GGAATTTTAAAGCGATCTGTCACCGCACTTTGCCATATTTAAACCTGATATC 1031
|||||
Db 1002 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 943
QY 1032 TTATTAAGTTAATAGATATATCCGTAGATTAATTAAGTATGTTAAACAGAGTAAACAA 1091

Db 942 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 883
QY 1092 ATATCTTAATATTTTAATTTCTGATTTATTTTGACGATATTTATTAATTTTAAGAA 1151
|||||
Db 882 TAAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 823
QY 1152 TTATCTATTTAGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1211
|||||
Db 822 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 763
QY 1212 AAAAGCATTAGAAATGAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 1271
|||||
Db 762 AAATACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 703
QY 1272 ACAATCTGTCATTTCTTGTATGATTAATTAATTAATTAATTAATTAATTAATTA 1331
|||||
Db 702 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 643
QY 1332 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1391
|||||
Db 642 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 584
QY 1392 CGGCTGGCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1451
|||||
Db 583 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 524
QY 1452 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1511
|||||
Db 523 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 464
QY 1512 TTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1571
|||||
Db 463 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 404
QY 1572 TGGAGAAATTAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1631
|||||
Db 403 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 344
QY 1632 GATATTAAGAAAGTCACTCTTTAAATTAATTAATTAATTAATTAATTAATTAAT 1691
|||||
Db 343 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 284
QY 1692 GAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1746
|||||
Db 283 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 224
QY 1747 AAATCTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1806
|||||
Db 223 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 164
QY 1807 CCTATGCTTTCCGATTTAAAAACAAGCAAAAAGATC 1844
|||||
Db 163 CATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 126

RESULT 3
AX251345/c
LOCUS AX251345 6286 bp DNA linear PAT 05-OCT-2001
DEFINITION Sequence 313 from Patent WO0168912.
ACCESSION AX251345
VERSION AX251345.1 GI:15984768

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
ARTIFICIAL SEQUENCES

REFERENCE
1 (bases 1 to 6286)
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with tumor suppressor genes and
oncogenes

JOURNAL Patent: WO 0168912-A 313 20-SEP-2001;
EpiGenomics AG (DE)

FEATURES
Location/Qualifiers

[illegible]

DEFINITION	Homo sapiens 3 BAC RP11-384L1 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION	AC069435
KEYWORDS	AC069435.19 GI:20340410
SOURCE	HTG.
ORGANISM	Homo sapiens.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 11861) Muzny,D.M., Adams,C., Adio-Ondola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbarta,J., Benton,J., Blinage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowe,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Bunyah,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Dudbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J., H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisgeed,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Matinez,E., Massey,E., Mahlhney,E., Mcleod,M.P., Meador,K., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunou,G., Ogranuye,N., Oviyedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Plimus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojoubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sison,I., Sodergren,E., Sonalle,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlecezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 11861)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (30-MY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 11861)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (30-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Apr 30, 2002 this sequence version replaced gii:20066152. INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones.

On Apr 30, 2002 this sequence version replaced gi:20066162
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
 gc-help@bcm.tmc.edu

Submitted (30-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Worley, K. C.
Direct Submission

Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 11861)

Submitted (30 May 2000)
human genome sequencing center, Baylor College of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Submitted (30-MAY-2000) Human Genome Sequencing Center, Department

2 (bases 1 to 111861)
Mortley, K.C.

Direct Submission
Unpublished

Wu, Y. F., Zhou, J., Zorrilla, S., Naylor, S. L., Weinstock, G. and Gibbs, R.

waiten, R., Washington, C., Watlington, S., Williams, G.,
Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,

Verger, J., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,

Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,
Tavor, T., Telford B., Thomas N., Thomas S., Ilmanen, K., Vasquez I.

Scott, G., Shën, H., Shooshtari, N., Sisson, I., Södergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,

Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M.,
Rojas, A., Rotjokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S.,

Nickerson, E., Nwokenkwo, S., Ogun, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L.,

Moore, S., Morgan, M., Moorish, I., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.

Ma, L., & Metzger, M. (2000). The effects of social support on the well-being of adolescents with chronic illness. *Journal of Pediatric Psychology, 25*, 101-110.

Mei, G., Metzger, M., Miner, G., Mitchell, T., Mohabbat, K., Moore, S., Morrison, M., Morris, C., Nelson, V., Neal, D., Neer, J., & Neuman, J. (2000). The effects of social support on the well-being of adolescents with chronic illness. *Journal of Pediatric Psychology, 25*, 101-110.

Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martínez, F., Massey, F., Mawhinney, F., McLeod, M. P., Meador, M.

Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Loulsead, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R.,

Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,

Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C.,

Gaddis, A., Gao, J., Garcia, A., Garner, I., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabriel, A., Gao, T., Garcia, A., Gerner, E., Gerra, N., Gill, P., Lamm, C., Logan, D., Lowary, C., Linn, J., C., Lubbock, N.,

Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Earnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M.,

Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A., Delaney, K. R., Delgado, O., Penn, A. L., Ding, Y., Dinh, H. H.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R.,

Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,

Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P., Barbarella, J., Benton, J., Bimlage, K., Blankenburg, K., Bonnin, D.,

Maznyk, D.M., Adams, C., Aalto, C.A., D., Ali, Osman, F.N., Alchuk, C.,
Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaric, I., Benton, I., Dimarco, P., Blankenburg, V., Bonita, D.

1 (bases 1 to 111861)
Muzny, D. M. , Adams, C. , Adio-Odiroja, B. , Ali-osman, F. R. , Allen, C. ,

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Homo sapiens.
Homo sapiens

AC089435.19 GI:20340410
HTG.

bnc library) complete sequence.
 AC069435
 2006425 10 ct. 20340410

Homo sapiens 3 BAC RP11-384L1 (Roswell Park Cancer Institute Human BAC library) complete sequence

Page

[illegible]

Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Bazzone, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferrier, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Idagawa, C., Jalili, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., Moshrefi, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunco, J., Pacle, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Plitman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirska, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M., and Venter, J.C.

Sequencing of *Drosophila* chromosome 3R, region 82E-82F

Unpublished

TITLE
JOURNAL
AUTHORS
REFERENCE

2 (bases 1 to 171105)

Celik, S.E., Abbayani, A., Arcana, T.T., Baxter, E., Blazer, R.G., Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L., Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Katta, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomutan, M.A., Mazda, P., Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacle, J.M., Park, S., Pfeiffer, B., Poon, L., Sequiera, A., Sethi, H., Sitr, E., Svirska, R.R., Man, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and Rubin, G.M.

Direct Submission

Submitted (12-MAY-1999) *Drosophila* Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Aug 24, 2001 this sequence version replaced gi:1683713.

Sequence submitted by:

Berkeley *Drosophila* Genome Project

Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to dbget@fruitfly.berkeley.edu.

FEATURES

source

1..171105

/organism="Drosophila melanogaster"

/strain="y: cn bw. sp"

/db_xref="taxon:7227"

/chromosome="3R"

/map="82E-82F"

/clone="BACR01D10 (D670)"

/clone.lib="PC1-98 (Roswell Park Cancer Institute

Drosophila melanogaster BAC library, partial EcoRI in

pBAC3.6)"

BASE COUNT

49208 a 36998 c 36264 g 48635 t

ORIGIN

Query Match 6.1%; Score 112.6; DB 3; Length 171105;

Best Local Similarity 48.2%; Pred. No. 1e-08;

Matches 407; Conservative 0; Mismatches 434; Indels 4; Gaps 3;

910 TTATACCAAGCAAAACAAAATAAACAAGAAAAATTTTCGAGGAAAAATAT 969

168200 TCATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 168259

970 TTGGAAATTTTAAAGGCGATCTGCTACGCGACTTTTGCATATTTAAAC 1029

168260 AAGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 168319

1030 TCTTATAGTAATAGATATATCGTTAGATTATATAGTATTAATAACAGTA 1089

168320 AATGAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 168379

1090 CAAT--AACTTATATATTTATTCGAATTTATTTGACGATGATTATTAAT 1147

168380 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 168439

1148 GAGATATATCTATATAGCTTAATAATAAATAAAGAGTAATATATGATGT 1207

DB 168440 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 168499

QY 1208 TTAATAAAGACCTAGAGAAATGAATAGAACATTATTAATAAAGCGTGATCAATCTT 1267

DB 168500 TGCTAATAATAGTAATAATAGTAATAATAAATAAATAAATAAATA 168558

QY 1268 TCTTACAAATCTTGGATTAATCTTGGATTAATGAATGAATGAATGCGG 1327

DB 168559 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 168618

QY 1328 ATATTTATTAACCTCAATTTATTCATTGGAAGATTGATCTTCTGATGTTGAC 1387

DB 168619 TAATGATTAATTCACATTAATAAATAAATAAATAAATAAATAAATA 168678

QY 1388 ATTTGGCGCTGGCTTCAATAAATAAATAAATAAATAAATAAATAAATA 1447

DB 168679 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 168738

QY 1448 AAATGAAATATGATTTTAAATAAATAAATAAATAAATAAATAAATAAATA 1507

DB 168739 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 168798

QY 1508 TTAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1567

DB 168799 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 168858

QY 1568 CATATGAGAAATTAATCAATGATGATGATGATGATGATGATGATGATGAT 1627

DB 168859 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 169318

QY 1628 AAAAGATATAAAGAAAGAGTACCTCTTAAATAATCTGATGATGATGATGAT 1686

DB 168919 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 169378

QY 1687 TAGATGAAATATTTATGATGAGAGGATTTCCAACTAAGATATTGAAATATG 1746

DB 168979 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 169038

QY 1747 AAAAT 1751

DB 169039 TAAAT 169043

RESULT 9

AC009212/c

LOCUS 190801 bp DNA linear INV 24-AUG-2001

DEFINITION *Drosophila* melanogaster, chromosome 3R, region 82E-82F, BAC clone

ACCESSION BACR01A18, complete sequence.

VERSION AC009212

KEYWORDS AC009212.5 GI:15284179

SOURCE *Drosophila melanogaster*.

ORGANISM *Drosophila melanogaster*

REFERENCE

1 (bases 1 to 190801)

Celik, S.E., Adams, M.D., Krommiller, B., Tyler, D., Man, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., An, H., Baldwin, D., Bazzone, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champe, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferrier, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Idagawa, C., Jalili, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., Moshrefi, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunco, J., Pacle, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Plitman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirska, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M., and Venter, J.C.

TITLE
JOURNAL
Unpublished

REFERENCE 2 (bases 1 to 190801)
 Authors Celniker, S.E., Agbayani, A., Arcane, T.T., Baxler, E., Blazek, R.G., Burenhoff, C., Champ, M., Chavez, C., Chew, M., Ciesiolka, L., Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L., Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Keaney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomonte, M.A., Macda, P., Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B., Poon, L., Segueira, A., Sethi, H., Sitr, E., Sylrskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zierian, L.L. and Rubin, G.M.
 Title Direct Submission
 Journal Submitted (06-AUG-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA
 Comment On Aug 24, 2001 this sequence version replaced gi:6957960. Sequence submitted by: Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory, MS 64-121 Berkeley, CA 94720
 This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.
 Location/Qualifiers
 source 1. 190801
 /organism="Drosophila melanogaster"
 /strain="y: cn bw sp"
 /db_xref="taxon:7227"
 /chromosome="3R"
 /map="82E-82F"
 /clone="BACR01A18 (D669)"
 /clone.lib="RPCI-98 (Roswell) Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in PBACs3.6")

BASE COUNT 53774 a 39988 c 41097 g 55942 t
 ORIGIN

Query Match 6.1%; Score 112.6; DB 3; Length 190801;
 Best Local Similarity 48.2%; Pred. No. 9.9e-09;
 Matches 407; Conservative 0; Mismatches 434; Indels 4; Gaps 3;

OY 910 TTATACCAAGCAAAACAAATAAATAACAGAAATTTTCGACGCAAAATATAT 969
 DB 180257 TCATCAAAAATTAATAATAATAATAATAATAATAATAATAATAATA 180198
 OY 970 TTGCAATTTTAAAGCGATACCTGACCGACCTTTGCCATATTTAAACCTGACTA 1029
 DB 180197 AAGTAAATTAATAATAATAATAATAATAATAATAATAATAATAATA 180138
 OY 1030 TCTTTAATAGTAATAGATATATCGGTAGATTAATAAGTATGTTAAACGAGTAATAA 1089
 DB 180137 AATGAACAATAATAATAATAATAATAATAATAATAATAATAATAATA 180078
 OY 1090 CAAT--AACTATATATTTAATTCGAATTAATTTGACAGATTTAATTAATTA 1147
 DB 180077 AAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 180018
 OY 1148 GAGATATATCTATTAGTTAATAATAATAATAATAATAATAATAATAATAATA 1207
 DB 180017 AAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 179958
 OY 1208 TTAATAAAGCATTAAGAAATGAAATAGAAATTAATAATAAAGCGGTATCAATCTT 1267
 DB 179957 TAGTAATAATAGTAATAATAATAATAATAATAATAATAATAATAATA 179899
 OY 1268 TCTTCAATACCTGATTAATTTGATAGATTAAGATTAATAAGTGAATGTGGCG 1327
 DB 179898 AAAATTAATAATAATAATAAGAAACATTAATAATAATAATAATAATAATA 179839
 OY 1338 ATATTATTAACCTCAATTTATCCATGATGAAGATTTGATCTTCTTGATGTGAGACA 1387
 DB 179838 TAATGATTAATTCATTAATAATAAGAAATTAATAATAATAATAATAATA 179779

OY 1388 ATTTGCTGGGCTTTCATATAATAATACAGACGAATGAATTTTATCAGCTA 1447
 DB 179778 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 179719
 OY 1448 AAACGAAATGATTTTAAATAATACGCTTTTATTCATTCGATTCATTCGAGAAAT 1507
 DB 179718 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 179659
 OY 1508 TTAATTCAGATGAGATGCTGATCAATACATCAAGAAATGATTAAGATTAATTAACA 1567
 DB 179658 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 179599
 OY 1568 CATATGGAATTAACCATGATTAATAATAATAATAATAATAATAATAATAATAATA 1627
 DB 179598 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 179539
 OY 1628 AAACATTAATAAAGAACTCAGCTCTTAAATAATCTCATGATGATGTTTAA-AG 1686
 DB 179538 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 179479
 OY 1687 TAGATGAATAATTTTATGATGAAGAGATTGCAACTTAAGATTTGCAATATGCTATG 1746
 DB 179478 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 179419
 OY 1747 AAAAT 1751
 DB 179418 TAAAT 179414

RESULT 10
 AX347260/c 12237 bp DNA linear PAT 01-FEB-2002
 LOCUS AX347260
 DEFINITION Sequence 2331 from Patent WO0200928.
 ACCESSION AX347260
 VERSION AX347260.1 GI:18495148
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM artificial sequences.
 REFERENCE 1
 Authors Olek, A., Piepenbrock, C., and Berlin, K.
 Title Diagnosis of diseases associated with the immune system
 Journal Patent: WO 0200928-A 2331 03-JAN-2002;
 Epigenomics AG (DE)
 FEATURES
 source Location/Qualifiers
 1. 12237
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 3050 a 207 c 3024 g 5956 t
 ORIGIN

Query Match 6.0%; Score 111.6; DB 6; Length 12237;
 Best Local Similarity 48.5%; Pred. No. 2.5e-08;
 Matches 364; Conservative 0; Mismatches 384; Indels 2; Gaps 2;

OY 1014 ATTTAAACCTGACTATCTTTATTAAGTTAATAGATATATCCGTAGATTAAGTATGT 1073
 DB 6912 AATCAACAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 6853
 OY 1074 TAAACAGAGTAATAAACAATATATATATTTATTCGATTTATTTGACAGCTGATT 1133
 DB 6852 AATCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 6793
 OY 1134 ATTTAATATATTAAGAGATATATCTATAGCTTA-ATATAACTTAATAAAGAGCTAAT 1192
 DB 6792 AATAATTAATCAATAATAATAATAATAATAATAATAATAATAATAATAATA 6733
 OY 1193 ATATGATTTGCTATTTAAATAAAGCATTAAGAAATGAATATGAACATTAATAAAGAGC 1252
 DB 6732 AAATTAATAAATAAATAAATAAATAATAATAATAATAATAATAATAATAATA 6674

QY 1253 GTGATATCAATCTTCTTACAACTGCTTCATTTGATATATGATTAAGCATTAATG 1312
 Db 6673 TAAAT 6614
 QY 1313 GTGATGATGTCGGGAT 1312
 Db 6613 ATAAAT 6554
 QY 1373 TTGATGTTGACACATATTCGGCTGGCTTTCATATATATATATATATATATATATAT 1432
 Db 6553 ATAAAT 6494
 QY 1433 ATTATATATGCTGCTAAATATATATATATATATATATATATATATATATATATAT 1492
 Db 6493 AAAAT 6434
 QY 1493 TCAATTCGACAT 1552
 Db 6433 AAAAT 6374
 QY 1553 GATTAGAT 1612
 Db 6373 AAAAT 6314
 QY 1613 TGATTCGAT 1672
 Db 6313 AAAAT 6254
 QY 1673 CGATGTTTAAAGTAT 1732
 Db 6253 AAAAT 6194
 QY 1733 GAAAT 1762
 Db 6193 AATTAACAATTAAT 6164

RESULT 11
 AE003603/c 294914 bp DNA linear INV 04-OCT-2000
 LOCUS Drosophila melanogaster genomic scaffold 14200013386043 section 5
 ACCESSION AE003603
 VERSION AE003603.2 GI:10727164
 KEYWORDS HTG.
 SOURCE Drosophila melanogaster.
 ORGANISM Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 294914)
 Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
 Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,J.R.F.,
 George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
 Sutton,G.G., Morten,J.R., Vandeil,M.D., Zhang,Q., Chen,L.X.,
 Brandon,R.C., Rogers,Y.H., Blazet,R.G., Champ,M., Pfeiffer,B.D.,
 Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gaber
 Miklos,G.L., Adill,J.F., Agbayani,A., An,H.J.,
 Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
 Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
 Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkov,D.,
 Botchan,M.R., Bouck,J., Brokstein,P., Brothier,P., Burtis,K.C.,
 Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
 Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
 Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
 Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
 Dunn,P., Durbin,K.J., Evangelista,C.C., Ferrer,C., Ferrier,S.,
 Fleischmann,W., Fosler,C., Gabriellian,A.E., Garg,N.S.,
 Gelibert,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
 Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
 Hernandez,J.R., Houck,J., Hostlin,D., Houston,K.A., Howland,T.J.,
 Wei,M.H., Idegawa,C., Jaisl,M., Kalush,F., Karpen,G.H., Ke,Z.,
 Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 SOURCE

2 (bases 1 to 294914)
 Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
 Direct Submission
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 On Oct 9, 2000 this sequence version replaced gi:7296751.
 Location/Qualifiers
 1..294914
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /chromosome="3R"
 complement(<8081..>8296)
 /gene="CG14663"
 /evidence=not_experimental
 /map="82F1-82F1"
 /db_xref="FLYBASE:FBan0014663"
 /db_xref="FLYBASE:FBgn0040683"
 complement(<8081..>8296)
 /gene="CG14663"
 /product="C134443"
 /evidence=not_experimental
 /db_xref="FLYBASE:FBan0014663"
 /db_xref="FLYBASE:FBgn0040683"
 complement(8081..8296)
 /gene="CG14663"
 /note="CG14663 gene product"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="AA52070.1"
 /db_xref="GI:7296792"
 /db_xref="FLYBASE:FBan0014663"
 /db_xref="FLYBASE:FBgn0040683"
 /translat="MPKRLPAARNASMLVPCLLPLPARMAPRPPGVPPAPRSICR
 ODPLBEGICNMSHLCGELGLVMD"
 <13958..>14959
 /gene="CG12590"
 /evidence=not_experimental
 /map="82F1-82F1"
 /db_xref="FLYBASE:FBan0012590"
 /db_xref="FLYBASE:FBgn0037294"
 <13958..>14959
 /gene="CG12590"
 /product="C134442"
 /evidence=not_experimental
 /db_xref="FLYBASE:FBan0012590"
 /db_xref="FLYBASE:FBgn0037294"
 13958..14959
 /gene="CG12590"
 /note="CG12590 gene product"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="AA52069.1"
 /db_xref="GI:7296791"


```

OY      1639 AAGAAAGTACGCTTTAAATATCGATTGATCGATGTTTAAAGATGAAAT 1698
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3952 ATAAATTAATAAGTAAATTAATAATAATAAACAACAAAAATGTTGTAAT 3893
OY      1699 TAAATGATGAGAGATTTGCAACTAAGATATGAAATATCGATGAAATCCATATG 1758
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3892 AATGATACGATACATTTAATACATTCACATTTAATAGATTTTATGAAACCAATGC 3833
OY      1759 ATG 1761
        ||
DB      3832 TTG 3830

RESULT 12
PEMALP3 67970 bp DNA linear INV 15-DEC-1999
LOCUS Plasmodium falciparum MALP3, complete sequence.
DEFINITION AL031746
ACCESSION AL031746
VERSION 1 GI:6594243
KEYWORDS HTG.
SOURCE Plasmodium falciparum 3D7.
ORGANISM Plasmodium falciparum 3D7.
REFERENCE 1 (bases 1 to 67970)
AUTHORS Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quall,M.
          and Barrell,B.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-1998) P.falciparum Genome Sequencing Consortium,
          The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
          CB10 1SA, UK

-COMMENT
On Dec 16, 1999 this sequence version replaced gi:5763807.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum. IMPORTANT: This
sequence is unfinished and does not necessarily represent the
correct sequence. Work on the sequence is in progress and the
release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc.

FEATURES
source          1..67970
                  Location/Qualifiers
                    /organism="Plasmodium falciparum 3D7"
                    /strain="3D7"
                    /db_xref="taxon:36329"
                    /chromosome="1"
                    /complement(join(1748..2598,2748..2848,2990..3276))
                    /gene="MALP3.01"
                    /complement(join(1748..2598,2748..2848,2990..3276))
                    /gene="MALP3.01"
                    /note="MALP3.01, conserved hypothetical protein, len: 412
aa, similarity: UPP0006 family eg to
YBL055C/YBL0512/YBL0511, YBP5_YEAST (418 aa), fasta
scores: opt: 316, E(): 1.1e-12, (33.2% identity in 271 aa
overlap)"
                    /codon_start=1
                    /product="conserved hypothetical protein, UPP0006 family"
                    /protein_id="CAB63556.1"
                    /db_xref="GI:6594244"
                    /translation="MKLVFHYIKYINLVFISIIPLKNSLKTYNDLRYSIVNKKYK
LOIKRSLNKKNNHNRKMEDESSFDIGSNLTDMPGVNKKKHENDLONVNRK
NNYNDKIITCTCAEIDKSLKICTYDPECKFLYLSGVPNTNCEYETIDKKEHEKE
ITAKKEYEFTKRFNEDVNSKMGKIKIDGKEDKNNLNLELLEKNDLTIPGEKY
NEKDEYELNKKIKIKYPNRIIVCGELGDFDLKFCISKYIQRKFFDLQVOMFN
LPMFLHMRNCSETEFKIVDIKFLPEKNGGVHSTFDKEDIHIIIVQYKNLIGVNG
CSLSKENINAVKRIPLNLILLETDAPMCVAKTHASVEYIKDYEKRAYNLKKIKN
IIKDDNITFEKRNPPYNTIA"
misc_feature    complement(2599..2610)
                  /gene="MALP3.01"
                  /note="potential splice acceptor sequence"
misc_feature    complement(2742..2747)
                  /gene="MALP3.01"
                  /note="potential splice donor sequence, atg/gttaa"
misc_feature    complement(2849..2861)

```

```

misc_feature    /gene="MALP3.01"
                  /note="potential splice acceptor sequence"
                  complement(2984..2989)
misc_feature    /gene="MALP3.01"
                  /note="potential splice donor sequence, aaa/gtaaa"
gene            5005..5496
                  /gene="MALP3.02"
                  /note="MALP3.02, hypothetical protein, len: 163 aa,
                  contains possible signal sequence"
CDS             5005..5496
                  /gene="MALP3.02"
                  /note="MALP3.02, hypothetical protein, MALP3.02"
                  /product="hypothetical protein, MALP3.02"
                  /codon_start=1
                  /protein_id="CAB63557.1"
                  /db_xref="GI:6594245"
                  /translation="MKLNNRFRVVCPIILIFFPLNSVVLGNNNNNNINPHETENAAK
AMRRLKGEISIKLDNGDELKIKDKERKHDSTKDWKSYSPISLSLEEKYSQDPLR
KKQETENANTKIIEDROEFYILNNDIEINATRFVLENNDELYIQSKSLIDIOS
LNN"
                  8020..10389
                  /note="possible cen1, region of very high [A+T] content"
                  14884..20352
                  /gene="MALP3.03"
                  14884..20352
                  /gene="MALP3.03"
                  /note="MALP3.03, putative ABC transporter, len: 1822 aa"
                  /codon_start=1
                  /product="putative ABC transporter"
                  /protein_id="CAB63558.1"
                  /db_xref="GI:6594246"
                  /translation="WTYKEVNGISNKKKKKKSCONISFLNLSFDMIRPLINDLIK
GDIELPNICRNFDPVYASKLEENLRDIEVDSFEYKSSNMFVYAHONSDASE
KKVYVYVYHNILMSLKTKFRFRIILISFYILETLVLVIGKFTDYVRRIIEGQIPY
YISFLKDFKVSGLVVMVINEFHLFEFLHFEFLINLKSVMFLYLYKICSN
NHLQMDAFYNTYRKESSTDEISDFDLISGKMASSSGIKRNNKNNKVEN
DYLINEPSTKMGDSNBRSDNPNVNTVIMSDVSPFPTSCNINLNVYKIF
MSFYFHRTKISNSGCIATWLSIALYSMLLPEPLPSLFPKSKYLIRKRDNNHVL
KEFLIKFENMSEFAKTYINIFRMKEMVCKIRLVLNIGVFISISDIYEVEYFLI
YLKRLNKEEIKETSIIMPLYVKKILISYVANNPNVNNVMEGIVNRLNLYINDH
LYNDIKNYFMVRYRNYEDYNIYVDKTLQVNDNTYKICFOEHKSNSTYANNKSHIE
MEKYEFYHKNNYHKNILNKOILSGLTNVDQNTYKICFOEHKSNSTYANNKSHIE
KKEEYENHTNSNSMTNMEFEKKEKKNNTYIILKLENSGSLSYDNKCONDHTLKNEN
LKRNLSIITIGVSGKSAFPHSLIGDNRNTHGNLIYINFPKKPILYVPONSLPMG
NIRSMILFGENEPLIYKYTLQSLDLNDLSTIEGDKYLIINDHNSLKGKVRCLLA
BALVEHYHIMHKLCTDYERKLIQPMELIKDILNKNLSYNNKSKLVANNIPEN
YLOKCLMDNNFYLYLDIDFTSLDPSISKIFLWLFCEKNISFKDKCSFLISNKS
TLDNFLIEDILDNOYEVNIIEIDKTLKYNGISSEYKNNLNLTKESHGSLNLT
IDYRIKLFDEVELNHYVKSNNKMYKEAYFYKGNTESEYFEDISNKEYTEKKKKKY
KKEHNKNNKNNNNNNNNNDHINIMNNHRYNDINLGPSTDSPTVYSLGNE
YTLDTYSSNSDKETIYKPLKDYHEEKNSSMPEYVSSNMATNPSNRYEDSSS
FKGSISLETYLYMFOQGVFLTQSVIEMLSIFTEDEKFEVLMMLSIISNNKESHD
TILQKQVLYLYMFOQGVFLTQSVIEMLSIFTEDEKFEVLMMLSIISNNKESHD
TYNNKGNVINEFIDISAFDYGFLRIYKAFIEFRCILSLIILYMRDCLFIEPY
KILYVGVNTGYIDIEEPLKNNINMAYLNNKIGIVSGKADSKRYKGFENYVYSYK
KKIPLVNGTYIDIEEPLKNNINMAYLNNKIGIVSGKADSKRYKGFENYVYSYK
KIYVGVNTGYIDIEEPLKNNINMAYLNNKIGIVSGKADSKRYKGFENYVYSYK
GINTGKNDLYIYNNKODKSNYKTIQSYKINSONDTILLTNDJCIRYSLVRLYN
RHKKRIILIDIEPIPLNLSNHYDELNSLTKAKSPNIIINHPNNTVLIISHANT
LSCDYLIVLKEKGETYRCSEYDVKTQSELSHLEMD"
rRNA            23896..31533
                  /gene="rRNA"
                  23896..31533
                  /gene="rRNA"
                  /note="region containing small subunit, 5.8S and large
                  subunit rRNA genes and spacer regions"
gene            complement(join(31966..32476,32675..32775))
                  /gene="MALP3.04"
                  complement(join(31966..32476,32675..32775))
                  /gene="MALP3.04"
                  /note="MALP3.04, conserved hypothetical membrane protein,

```



```

len: 203 aa, similarity: P. falciparum chromosome 2,
PF01010W, 096126 predicted integral membrane protein (255
aa), fasta scores: opt: 335, E(): 4.9e-15, (36.1% identity
in 191 aa overlap)"
/codon_start=1
/product="conserved hypothetical membrane protein,
MALP3.04"
/protein_id="CAB63559.1"
/db_xref="GI:6594247"
/translation="MKSKYFTINWTLLEFLTLFLTYNYNDFSKTKFRNNIKIIN
KRRITAESEEEKYPMEDFCILNEBELIRPHNDSPLEPHYENIDKINLSINS
TKIMKFTIKMRQNYERKTDNNMNNRDMHYKMANIYLKYHKLINTLKDTNPI
HDKERTTIWKIKIOEDIEYFLFNLOVEMLRILTLFLFYKNKE"
/complement(32477..32486)
/gene="MALP3.04"
/feature="potential splice acceptor sequence"
/complement(32669..32674)
/gene="MALP3.04"
/feature="potential splice donor sequence, aag/gtataa"
/join(36657..36743,36864..37343)
/gene="MALP3.05"
/join(36657..36743,36864..37343)
/gene="MALP3.05"
/codon_start=1
/product="hypothetical protein, MALP3.05"
/protein_id="CAB63560.1"
/db_xref="GI:6594248"
/translation="MKIKMNSGIFETKILICISFICEFECGKMTSYRDLWYEN
CENYSIDRLAESESESKETKVDIPNIELKSLNTEYEMKELIVSFMNNMIN
IANEVLKNIHSFENIENISFLINDSSKSPKFLKEFGSIFPHMLNNPKLLFDLCO
RNPILHICILVLAIVYENKNEFC"
/36744..36749
/gene="MALP3.05"
/feature="potential splice donor sequence, aag/gtataa"
/36854..36863
/gene="MALP3.05"
/feature="potential splice acceptor sequence"
/complement(38049..39995,40210..40284))
/gene="garp"
/complement(join(38049..39995,40210..40284))
/gene="garp"
/feature="garp"
/feature="MALP3.06, garp, len: 673 aa, similarity: almost
identical to GARP PLAF (678 aa), fasta scores: 97.6%
identity in 678 aa overlap"
/codon_start=1
/product="hypothetical garp protein"
/protein_id="CAB63561.1"
/db_xref="GI:6594249"
/translation="MNVLFSYNICILFFVCTLNSTKCFSGNLKLNLIKNSFDS
ITGRLENELEKNDKDNSEKTLKEEDKDPPTSNDKLNKNNNEISSTDP
TNIINNDKDNSESVKDKKDKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
DKKKNSESVSLYKQHKPKNATHEGENLYEEMVSENNNAOGGLLSSPYRE
OGGCGIISVHETSNDKNDKENDSEKEDHQOEMDKTLDKRRKOLSEKMOE
KLEKKKKQOEKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
HKENETMQOPOISETNEINMLPLPLDVTPTPEEKKEEHEEHEEHEG
KKEHEHKEHEHKEHEHKEHKKKKKKKKKKKKKKKKKKKKKKKKKKKK
DKDGYIINLEKKEACEPHITVESRPLSQPOCKLIDEBOLILMKSKYEEKNLSIO
EOLIGITIGRYNVPRKDNHKKMAKLEELQKQKQKQKQKQKQKQKQKQKQ
EDDEEVEEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
DDDAEEDDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
complement(39996..40005)
/gene="garp"
/feature="potential splice acceptor sequence"
/complement(40204..40209)
/gene="garp"
/feature="potential splice donor sequence, aag/gtataa"
/join(45401..46396,46562..50233)
/gene="MALP3.07"

```

Query Match 6.0%; Score 110; DB 3; Length 67970;
 Best Local Similarity 49.4%; Pred. No. 3.1e-08;
 Matches 433; Conservative 0; Mismatches 425; Indels 18; Gaps 5;

```

QY 883 AAAATCTAATCCAGGTTGGGTTTTTATACAGAGCAAAACAAAATAAACA 942
   || || || || || || || || || || || || || || || || || || ||
Db 8263 AATAATATATATATATATATATATATATATATATATATATATATATAT 8322
QY 943 AGAAAAATTTTCAGCGAAAAAATTTTGGAAATTTTTTAAAGCGATCTGACCGC 1002
   || || || || || || || || || || || || || || || || || || ||
Db 8323 A-TAAATATATATATATATATATATATATATATATATATATATATATAT 8381
QY 1003 ACTTTGCCATTTTAAACCTGACATCTTTTAAAGTTAATAGATATATCCGTAGATT 1062
   || || || || || || || || || || || || || || || || || || ||
Db 8382 ATTTTAAATATATATATATATATATATATATATATATATATATATATAT 8441
QY 1063 ATTAAGATGTATTAACGATTAACGATTAACGATTAACGATTAACGATTAACG 1122
   || || || || || || || || || || || || || || || || || || ||
Db 8442 AATTAATATATATATATATATATATATATATATATATATATATATATAT 8501
QY 1123 TGACAGTATATATATATATATATATATATATATATATATATATATATAT 1182
   || || || || || || || || || || || || || || || || || || ||
Db 8502 ATTAATATATATATATATATATATATATATATATATATATATATATAT 8561
QY 1183 AGAGGTAAATATATATATATATATATATATATATATATATATATATATAT 1242
   || || || || || || || || || || || || || || || || || || ||
Db 8562 ATATTAATATATATATATATATATATATATATATATATATATATATAT 8614
QY 1243 AAAAAGACGCGATATATATATATATATATATATATATATATATATATAT 1302
   || || || || || || || || || || || || || || || || || || ||
Db 8615 ATTAATATATATATATATATATATATATATATATATATATATATATAT 8674
QY 1303 GCATTAATATATATATATATATATATATATATATATATATATATATAT 1360
   || || || || || || || || || || || || || || || || || || ||
Db 8675 ATATTAATATATATATATATATATATATATATATATATATATATATAT 8734
QY 1361 GTTTTGAATCTTCTTGATGTTGACACATTTTGGCGCTTTCATATATATACACA 1420
   || || || || || || || || || || || || || || || || || || ||
Db 8735 AATTAATATATATATATATATATATATATATATATATATATATATAT 8794
QY 1421 GACGAATATATATATATATATATATATATATATATATATATATATATAT 1480
   || || || || || || || || || || || || || || || || || || ||
Db 8795 AATTAATATATATATATATATATATATATATATATATATATATATAT 8851
QY 1481 TTATTCATTCGATCAATTCGGAATTTAATATAGATGATGATGATGATGATG 1540
   || || || || || || || || || || || || || || || || || || ||
Db 8852 TATTAATATATATATATATATATATATATATATATATATATATATAT 8911
QY 1541 AATTCATTAAGATATATATATATATATATATATATATATATATATATAT 1600
   || || || || || || || || || || || || || || || || || || ||
Db 8912 ATATATATATATATATATATATATATATATATATATATATATATAT 8971
QY 1601 AAGAGATATATATATATATATATATATATATATATATATATATATAT 1660
   || || || || || || || || || || || || || || || || || || ||
Db 8972 AATATATATATATATATATATATATATATATATATATATATATATAT 9031
QY 1661 ATCTGATTCATGATGATTTTAAAGTAGATG-----AAAAATTTATGTAAGAGAT 1715
   || || || || || || || || || || || || || || || || || || ||
Db 9032 ATTCGCTTATATATATATATATATATATATATATATATATATATATAT 9091
QY 1716 TTCAACTAATATATATATATATATATATATATATATATATATATATAT 1751
   || || || || || || || || || || || || || || || || || || ||
Db 9092 TAATTAATATATATATATATATATATATATATATATATATATATAT 9127

```

RESULT 13

AE001398/c 14867 bp DNA linear INV 16-APR-2002
 LOCUS
 DEFINITION
 Plasmodium falciparum chromosome 2, section 35 of 73 of the
 complete sequence.
 ACCESSION
 AE001398 AE001362
 VERSION
 AE001398.1 GI:3845197
 KEYWORDS
 SOURCE
 ORGANISM
 Plasmodium falciparum 3D7.
 Plasmodium falciparum 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

[illegible][illegible]

COMMENT

Cambridgeshire, CB10 1SA, UK: E-mail enquiries: humbrey@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

On Feb 1, 2002 this sequence version replaced g1:1697351.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; SW: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Ch1
RP11-269F19 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-269F19. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true left end of clone RP5-88207 is at 88488 in this sequence. The true right end of clone RP4-678E16 is at 2000 in this sequence.

FEATURES

source

1. 90487

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="1"

/clone="RP11-269F19"

/clone_11b="RP11-11.1"

2097. 2223

/note="Sequence from reads from a short insert library derived from a clone PCR. Restriction digest data confirm the assembly."

misc-feature

BASE COUNT 22057 a 22328 c 23459 g 22643 t

ORIGIN

Query Match 5.9%; Score 108.8; DB 9; Length 90487;
Best local Similarity 48.4%; Pred. No. 4.5e-08;
Matches 449; Conservative 0; Mismatches 472; Indels 7; Gaps 5;

QY 922 CAACACAAAAATTAACACAGAAAAATTTGACGCAAAAAATTTTGAATTTT 981
DB 58868 CAAAAACAAAAACAAAAACACAGCTGTAGATCTGTGACACTGTGATTAAT 58927
QY 982 AAAGGAGATCTGCTACCGCACTTTGCCATTTTAAACGCACTATCTTATAAGT 1041
DB 58928 TAAGAGCCATTTAGGCGCAACTGTGCCCCCTCCCAAAATTCATATTAATATATA 58987
QY 1042 AATAGATATATCCGTAGATTATTAAGTATGTTAAACAGATTAACATATCTATA 1101
DB 58988 AATATATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATA 59047
QY 1102 TATTAATCTGATTAATTTGACAGTGATTAATTAATATTAATTAAGATATATCTAT 1161
DB 59048 TATTAATATTAATATTAATATTAATATTAATATTAATTAATTAATATTAATATA 59107
QY 1162 AGCTTAATATTAATTAACAGATTAATATTAATATTAATTAATTAATTAATTAATTA 1221
DB 59108 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 59167
QY 1222 GAAATGAAATGACATTAATAAAA-AGACGCTGATATCAATCTTTCTTACATATCT 1280

DB 59168 TATATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 59227
QY 1281 GCATTACTTGTATATGATTAACATTAATGTTGATGATGCGCAT-ATTATTAAC 1339
DB 59228 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 59287
QY 1340 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 59347
DB 59288 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 59347
QY 1400 CTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1459
DB 59348 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 59407
QY 1460 ATTTTAAAAATCTGCTTTTATTCATTCGTCATTCGTCGTCGTCGTCGTCGTCGTCGTCG 1519
DB 59408 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 59467
QY 1520 AGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1579
DB 59468 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 15926
QY 1580 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1639
DB 59527 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1699
QY 1640 AGAAGTACCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1699
DB 59586 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1759
QY 1700 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1759
DB 59646 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 17902
QY 1760 TGATCCAGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1819
DB 59703 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1819
QY 1820 CATTAAAAACACAGCAAAAGAGCT 1847
DB 59763 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1847

RESULT 15
AC103590 213056 bp DNA linear PRI 30-JAN-2002

LOCUS
DEFINITION Homo sapiens chromosome 1 clone RP11-544K18, complete sequence.

AC103590 AL357933
VERSION
AC103590.2 GI:18425292

KEYWORDS

HTG.

SOURCE

ORGANISM

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 213056)

Authors

Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,

Saenphimachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.

Direct Submission

Unpublished

2 (bases 1 to 213056)

Authors

Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.

Direct Submission

Submitted (29-NOV-2001) Genome Center, University of Washington,

Box 352145, Seattle, WA 98195, USA

3 (bases 1 to 213056)

Authors

Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,

Saenphimachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.

Direct Submission

Submitted (30-JAN-2002) Genome Center, University of Washington,

Box 352145, Seattle, WA 98195, USA

On Jan 30, 2002 this sequence version replaced g1:17149455.

Genome Center

Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgchgs@u.washington.edu
 Drafting Center: SC

----- Project Information

Center project name: chr-1
 Center clone name: RP11-544K18 (sc0176)

----- Summary Statistics

Sequencing vector: Plasmid; 6% of reads
 Sequencing vector: Plasmid; 108752; 94% of reads
 Chemistry: Dye-terminator; Big Dye; 33% of reads
 Chemistry: Dye-terminator; Big Dye; 33% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 212910 bases at least Q40
 Consensus quality: 212999 bases at least Q30
 Consensus quality: 213045 bases at least Q20
 Insert size: 212874; sum-of-contigs
 Quality coverage: 11.3x in Q20 bases; sum-of-contigs

----- Overlapping Sequences:

5': RP11-249G23 AL360008, 2000-bp overlap
 3': RP11-568G11 AL445435, 55927-bp overlap

----- Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

----- Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

----- EcoRI

----- HindIII

----- BglII

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
8696	9333	2083	2207	4498	4445
6	<800	6382	6525	2067	2067
8596	8488	512	<800	4570	4720
6075	6254	449	<800	8205	8225
744	803	3942	3884	6233	6198
1451	1430	33	<800	46	<800
5074	4975	4226	4220	6178	6198

8472	8488	2573	2547	1093	1117
791	803	2206	2207	1370	1329
1630	1634	2392	2547	580	<800
976	947	4969	4925	1805	1789
2759	2807	4673	4652	479	<800
310	<800	1874	1855	19	<800
4237	4205	1946	1960	3931	4030
3832	3797	3486	3458	3165	3213
3874	3907	197	<800	2530	2554
1078	1066	1107	1067	3866	3836
1295	1254	3787	3747	944	944
3092	3047	4066	4005	3013	3006
6272	6254	6168	6150	2222	2188
7581	7483	1075	1067	2324	2374
377	<800	1831	1855	704	<800
2328	2392	910	895	3306	3213
1985	1975	1344	1286	1282	1329
4079	4075	193	<800	7936	7908
9368	9333	405	<800	5478	5466
1675	1634	9929	10079	1144	1117
9377	9333	10756	10538	20101	20463
1996	2119	9627	9680	8397	8591
34994	34910	607	<800	8571	8591
309	<800	403	<800	4082	4030
838	803	10148	10538	5725	5466
2150	2119	10455	10538	5930	5737
6918	7078	6913	6973	635	<800
1374	1327	1677	1648	5211	5124
1770	1769	5476	5444	5529	5737
2610	2635	1075	1067	1393	1329
2246	2297	1434	1394	7324	7355
140	<800	595	<800	5699	5737
2470	2539	1075	1067	7123	7086
6	<800	17612	17124	3985	4030
2081	1975	18631	18193	603	<800
31	<800	5953	5976	2550	2554
5702	5672	2204	2207	1887	1914

THIS PAGE BLANK (USPTC)

[illegible]

QY	1152	TATATCTATTTAGCTTAATATTAAGCTAAAAAGAGTAAATATATGTGATGCTGTAATTAA	1211
Db	822	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	763
QY	1212	AAAAAGCATTAGAAAAATGAAATAGAACATTTATAAAAAAGACGGTGAATTCAAATCTTCTT	1271
Db	762	AAATAGCTTAATTAATTAACGAAATTAATAATTAATTAATTAATTAATTAATTAATTAAT	703
QY	1272	ACATACTCTGATTTACTTTGATTAATAGATTAAGCAATTAATTAATGCTATGAATGTGGCAATAT	1331
Db	702	AAATTAATTAATTAATAACAAATTAATAATTAATTAATTAATTAATTAATTAATTAAT	643
QY	1332	TATTAAGCTCAAAATTTATCCATTTGATGAAAGTTTGTGATCTCTGATGTGAGCAAAATTT	1391
Db	642	AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	584
QY	1392	CGGCTGGGCTTTCATTAATAATTAATACAGAGCAAGTAATGATATTTATATCAGCTAAAC	1451
Db	583	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	524
QY	1452	TGAAATATGATTTTAAAAAATTAATCTCGTCTTATTCATTCGATTCGATTCGATTCGATTCGAT	1511
Db	523	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	464
QY	1512	TTACGATGAGATTTGATACAAATACATCAAGAAATGATTAATTAATTAATTAATTAATTAAT	1571
Db	463	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	404
QY	1572	TGGAGAAATACCATATGATTAATTAATAAATGAAGATTAATTAATTAATTAATTAATTAAT	1631
Db	403	TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	344
QY	1632	GATTTAAATGAAGAACTACGCTCTTAAATAATCTGCATTTGATCGATTCGATTCGATTCGAT	1691
Db	343	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	284
QY	1692	GAAAAATTTTATTTGATGAAGAGATTTGCAACTTAAGATTAATGA-----AATATCGTATG	1746
Db	283	AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	224
QY	1747	AAAAATCCTATTTGATGATCCAGATGATGGCATTAAGAAAATCCAAATTCGACAGGAAGAAATG	1806
Db	223	AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	164
QY	1807	CCTATGCTTTCGCAATTAATAAAAAACAAGCAAAAAAGAG	1844
Db	163	CATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	126
RESULT 3			
AAS46591/c			
ID	AAS46591	standard; DNA; 6286 BP.	
XX	AAS46591;		
XX	18-DEC-2001	(first entry)	
XX		Tumour suppressor gene derived chemically modified sequence #313.	
XX		Human; tumour suppressor gene; oncogene; antitumour; cytostatic;	
XX		cancer; tumour; Cpg dinucleotide; single-nucleotide polymorphism; SNP;	
XX		cytosine methylation; ds.	
XX	Homo sapiens.		
XX	WO200168912-A2.		
XX	20-SEP-2001.		
XX	15-MAR-2001; 2001WO-EP02955.		
XX	15-MAR-2000; 2000DE-1013847.		
XX			

PR	06-APR-2000; 2000DE-1019058.
PR	07-APR-2000; 2000DE-1019173.
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
PA	(EPIG-) EPIDEMIOLOGY AG.
PI	Olek A., Piepenbrock C., Berlin K.
XX	WPI: 2001-602752/68.
XX	
PT	Fragments of chemically modified genes associated with tumour suppressor
PT	genes and oncogenes, useful in designing primers and probes for
PT	analysing diseases associated with cytosine methylation state e.g.
PT	cancer
XX	-
PS	Claim 1; SEQ ID NO 313; 27pp; English.
CC	The invention relates to a nucleic acid comprising a sequence of 18
CC	bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC	bisulphite, of genes associated with tumour suppression and
CC	oncogenes having a sequence taken from 536 (actually 533 since
CC	numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC	(SS) and sequences complementary to (SS). The nucleic acid may be a
CC	peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC	form part of a set of probes for detecting the cytosine methylation state
CC	and/or single nucleotide polymorphisms and also to be used in an
CC	array for analysing diseases associated with CpG dinucleotides e.g.
CC	cancers and tumours. The probes can also be used in a method for
CC	ascertaining genetic and/or epigenetic parameters for the diagnosis
CC	and/or therapy of existing diseases or the predisposition to specific
CC	diseases, by analysing cytosine methylations. The parameters may be
CC	compared to another set of genetic and/or epigenetic parameters, the
CC	differences serving as basis for diagnosis and/or prognosis events which
CC	are disadvantageous to patients. The present sequence is one of the
CC	533 genomic sequences derived from tumour suppressor genes and
CC	oncogenes.
CC	Note: The sequence data for this patent did not form part
CC	of the printed specification, but was obtained in electronic
CC	format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences.
CC	
XX	Sequence 6286 BP; 1508 A; 124 C; 1538 G; 3116 T; 0 other;
SO	
Query Match	6.7%; Score 123.2; DB 22; Length 6286;
Best Local Similarity	48.9%; Pred. No. 1.6e-11;
Matches 362; Conservative	0; Mismatches 373; Indels 5; Gaps 1
QY	1012 ATATTTAAACCTGACTCCTTCCTTTAATAGTATATAGATCATCGGTAGATTTAAAGTAT
DB	1012 ATATTTAAACCTGACTCCTTCCTTTAATAGTATATAGATCATCGGTAGATTTAAAGTAT
QY	4540 ATTAATATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
DB	4540 ATTAATATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
QY	1072 GTTAAAAACGATAAAACAATACTTATATTTAATTCGAATTTATATTGGACATGA
DB	1072 GTTAAAAACGATAAAACAATACTTATATTTAATTCGAATTTATATTGGACATGA
QY	4480 TAAATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
DB	4480 TAAATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
QY	1132 TTATTTAATATATTAAGAATATATCTTTTGCTTAATATATACGAAAAAAGAGCTAA
DB	1132 TTATTTAATATATTAAGAATATATCTTTTGCTTAATATATACGAAAAAAGAGCTAA
QY	4420 TAAATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
DB	4420 TAAATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
QY	1192 TATATGATTTCTGATTTTAAAAAGCATAGCAATTAAGAACATTTATAAAAAGAC
DB	1192 TATATGATTTCTGATTTTAAAAAGCATAGCAATTAAGAACATTTATAAAAAGAC
QY	4360 TAAATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
DB	4360 TAAATTAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
QY	1252 GGTCATATCAAACTCTTCTTCAATACCTTGCAATTCCTTGATATATGATTAAGCATTAAT
DB	1252 GGTCATATCAAACTCTTCTTCAATACCTTGCAATTCCTTGATATATGATTAAGCATTAAT
QY	1312 GGTATGATGATGTGGGATATATATAACCAATATATTCATTGATGAAGAAGTTTGATCTT
DB	1312 GGTATGATGATGTGGGATATATATAACCAATATATTCATTGATGAAGAAGTTTGATCTT
QY	4240 AAAATTAATTAAT-----AAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
DB	4240 AAAATTAATTAAT-----AAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
QY	1372 CTGATGCTTGAGCAATTTTCGGCTTCCTTCAATTAATTAATTAATTAATTAATTAATTA
DB	1372 CTGATGCTTGAGCAATTTTCGGCTTCCTTCAATTAATTAATTAATTAATTAATTAATTA

[illegible]

QY 1463 TTAATAAATACCTGCTTCTTATTCATTCGATCAATTGGAGAATTTATTTACGATGAGA 1522
 Db 4107 TAAATTTAAAAACATTAATCAATATACGAAACGCAAAACAAAAATCAATACCTACC 4048
 QY 1523 TGGGTCAATACATCAAGAAATGATTAAGATTAATTAATTTACATATATGGGAATTA 1582
 Db 4047 TATTCATCAATAAACCTCTAAAAACGAACTAATTAATTAATTAATTAATTAATTA 3988
 QY 1583 CCATATATTAATTAATTAATTAAGATATATTTGATTTTGAATTAACAAGATTTAAAGA 1642
 Db 3987 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3928
 QY 1643 AACTGACCTCTTAAAAATATCTGATTCATTCGATGATTTTAAAGTAGATCAAAATTTAT 1702
 Db 3927 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3868
 QY 1703 TGTATGAAGAGATTTTGCACACTAAAGATTTTGAAGAAATTCGTATGAAT 1751
 Db 3867 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3819

RESULT 7

ABL34358/C
 ID ABL34358 standard; DNA; 12237 BP.

AC ABL34358;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 2331.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KM antiarteriosclerotic; anti-HIV; anticonvulsant; ophthalmologic;
 KM neuroprotective; anti-inflammatory; antiproliferative;
 KM antihemorrhagic; antihemorrhagic; antidiabetic; antipsoriatic;
 KM antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KM acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KM gene; ds.

OS Homo sapiens.

PN WO200200928-A2.

XX 03-JAN-2002.

PD 02-JUL-2001; 2001MO-EP07537.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX
 PS Claim 1: SEQ ID NO 2331; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.

XX Sequence 12237 BP; 3050 A; 207 C; 3024 G; 5956 T; 0 other;

Query Match 6.0%; Score 111.6; DB 24; Length 12237;
 Best Local Similarity 48.5%; Pred. No. 1.2e-09;
 Matches 364; Conservative 0; Mismatches 384; Indels 2; Gaps 2;

QY 1014 ATTAAACCTGACATCTTATTAAGTATTAATGATATATCCCTGATATTAAGTATGT 1073
 Db 6912 AATCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6853
 QY 1074 TAAACGAGTAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1133
 Db 6852 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6793
 QY 1134 ATTATATATTAATTAAGATATATCTTATGCTTAA-ATATACATAAAAAAGAGTAAAT 1192
 Db 6792 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6733
 QY 1193 ATATGATGTTGTATTTAAAAAAGCATTTAGAAAATGAATAGACATTTAAAGAGCG 1252
 Db 6732 AATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 6674
 QY 1253 GTGATATCAATCTTCTTACATTAATGATTAATTAATTAATTAATTAATTAATTA 1312
 Db 6673 TAAATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 6614
 QY 1313 GTGATGAATGTCGCGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1372
 Db 6613 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6554
 QY 1373 TTGATGTTGAGCAATTTTCGGCTGCGCTTTCATTAATTAATTAATTAATTAATTA 1432
 Db 6553 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6494
 QY 1433 ATTATATCAAGTAAACGAAATGATTTTAAATAATTAATTAATTAATTAATTAATTA 1492
 Db 6493 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6434
 QY 1493 TCATTTGGAGAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1552
 Db 6433 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6374
 QY 1553 GATTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1612
 Db 6373 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6314
 QY 1613 TGAATTCGAATTAACCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1672
 Db 6313 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6254
 QY 1673 CGATGTTTAAAGTGAATGAATAATTTATTTGATGAAGAGATTTGCAACTAAAGATTT 1732
 Db 6253 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6194
 QY 1733 GAAATATATGATGAATAATTCCTATTTGATGA 1762
 Db 6193 AATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6164

RESULT 8
 AAK73165
 ID AAK73165 standard; DNA; 20420 BP.

XX AAK73165;

AC 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27977.

KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KM cytosine; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226881.
PR 22-AUG-2000; 2000US-0226886.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235844.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0244676.
PR 08-NOV-2000; 2000US-0244677.
PR 08-NOV-2000; 2000US-0244678.
PR 08-NOV-2000; 2000US-024523.
PR 08-NOV-2000; 2000US-024524.
PR 08-NOV-2000; 2000US-024525.
PR 08-NOV-2000; 2000US-024526.
PR 08-NOV-2000; 2000US-024527.
PR 08-NOV-2000; 2000US-024528.
PR 08-NOV-2000; 2000US-024532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
XX
PA Rosen CA, Barash SC, Ruben SM;
PI

```
XX WPI: 2001-483426/52.
DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
PS Disclosure: SEQ ID NO 27977; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins, and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 20420 BP: 4834 A; 5403 C; 5630 G; 4553 T; 0 other;
Query Match 5.3%; Score 98.2; DB 22; Length 20420;
Best Local Similarity 48.9%; Pred. No. 1.9e-07;
Matches 346; Conservative 0; Mismatches 358; Indels 3; Gaps 3;
OY 1032 TTTTAACTTAAATAGATATATCCCTAGATTTAAAGATGTTAAACGAGTAAACCA 1091
DB 4094 TATATATGTATATATATATATATATGAAAAATATATATATATATATATATGAAAA 4153
OY 1092 ATAACTTATATATATATATATATATATATATATATATATATATATATATATATAT 4151
DB 4154 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 4213
OY 1152 TATATCTATTTAGCT-TAAATATTAACATAAAAAAGAGTAAATATATGATGCTGATTTA 1210
DB 4214 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 4273
OY 1211 AAAAAAGCTTAGAAAAATGAAATAGACATTTAAAAAAGACGCGTATTCAAATCTTTCT 1270
DB 4274 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 4333
OY 1271 TACAATCTGTCAT-TACCTTGTATATGATTAAGATTAAGATGATGATGCGCAT 1329
DB 4334 ATAAATATATATATATGAAAAATATATATATATATATATATATATATATATAT 4393
OY 1330 ATTATAACTCAAAATTTTCCATTTGATGATGATGATGATGATGATGATGATGAT 1389
DB 4394 ATATATATATGAAAAATATATATATATATATATATATATATATATATATATAT 4453
OY 1390 TTCGGCTGGCTTCAATTAATATATACAGACCAAAATGATATTTATCATCAGCTAAA 1449
DB 4454 ATATATGAAAAATATATATATATATATATATATATATATATATATATATATAT 4512
OY 1450 ACTGAAATGATTTTAAAAAATACGCTTTTATTTCAATTCGATCAATTTGGAGAAATTT 1509
DB 4513 TATGAAATATATATATATATATATATATATATATATATATATATATATATATAT 4572
OY 1510 AATTATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1569
DB 4573 AAAATATATATATATATATATATATATATATATATATATATATATATATATAT 4632
OY 1570 TATGAGAAATATACATATGATATATATATATATATATATATATATATATATATAT 1629
DB 4633 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 4692
```

```
OY 1630 AAGATATATAAGAAAGTACGCTCTTAAATAATATGATGATGATGATGATGATGATGATGATGAT 1689
DB 4693 AATATATATATGAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 4752
OY 1690 ATGAAAAATTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1736
DB 4753 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4799
RESULT 9
ABK69933
ID ABK69933 standard; DNA; 20420 BP.
XX
AC ABK69933;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human secreted protein gene 68 genomic DNA fragment #21.
XX
KW Human; ds; secreted protein; gene therapy; immunosuppressive;
KW antirheumatic; antiproliferative; cytostatic; cardiant;
KW vasculoprotective; neuroprotective; autoimmune disease; neoplasm;
KW viricide; fungicide; opthalmological; hyperproliferative disorder; cardiac arrest;
KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; skin aging;
KW epithelial cell proliferation; food additive.
XX
OS Homo sapiens.
XX
PN WO200226931-A2.
XX
PD 04-APR-2002.
XX
PF 24-SEP-2001; 2001WO-US29871.
XX
PR 25-SEP-2000; 2000US-234925P.
XX
PI 12-JAN-2001; 2001WO-US00911.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;
PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
PI Endress GA, Mucenski M, Edner R;
XX
DR WPI: 2002-362489/39.
XX
PT Novel 71 isolated secreted polypeptides and polynucleotides encoding
PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis,
PT asthma -
XX
PS Example 2; Page 1440-1445; 1478pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (or its
CC fragment, homologue complement or allelic variant) encoding a human
CC secreted protein (and its fragment, domain, epitope, variant, secreted
CC form and species variant). Also included are a recombinant vector
CC comprising the nucleic acid, a recombinant host cell comprising the
CC vector, an antibody against the secreted protein, a recombinant host cell
CC that expresses the secreted protein and a method of identifying a binding
CC partner of the secreted protein. The nucleic acid and protein are used to
CC prevent, diagnose, treat or ameliorate a medical condition in e.g.
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
CC for example autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC ocular disorders e.g. corneal infection. Many other diseases and
CC disorders are listed in the specification. The polypeptides can also be
CC used to aid wound healing an epithelial cell proliferation, to prevent
```


CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. The present
CC sequence represents a ds DNA fragment of the gene for a novel human
CC secreted protein of the invention.

XX Sequence 20420 BP; 4834 A; 5403 C; 5630 G; 4553 T; 0 other;

Query Match 5.3%; Score 98.2; DB 24; Length 20420;

Best Local Similarity 48.9%; Pred. No. 1.9e-07;

Matches 346; Conservative 0; Mismatches 358; Indels 3; Gaps 3;

```
QY 1032 TTTATTAAGTAAATAGATATATCCGTTAGATTATTAAGTGTAAAAACGAGTAAAAACA 1091
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4094 TATATATGATATATATATATATGAAAAATATATATATATATATATATATGAAAAA 4153
QY 1092 ATACTATATATATATATTCGATTTATTTGACAGATATATTAATATTAACAGA 1151
    || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 4154 TATATATTAATATATATATGAAAAATATATATATATATATATATATATATATATA 4213
QY 1152 TATATCTATTAAGT-TAAATTAACAAAAAGAGTAATATATGATGTGATTTTA 1210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4214 TATATATATATATATATATATATATATATATATATATATATATATATATATATA 4273
QY 1211 AAAAAAGATTAGAAAAATGAAATAGACATTATTAATAAAAAACGATGATTCATCTTCT 1270
    || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 4274 ATATATATATATATATATATATGAAAAATATATATATGAAAAATATATATATATAT 4333
QY 1271 TACATATCTTCGAT-TACTTGTATATAGATTAACATTAAATGCGTATGATGCGCAT 1329
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4334 ATAAATATATATATATGAAAAATATATATATATATATATATATATATATATATAT 4393
QY 1330 ATTAATTAACCTCAAAATTTATCCATGATGAAAGTTTGTATCTTCTGATGTTGACACAA 1389
    || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 4394 ATATATATATGAAAAATATATATATATATATATATATATATATATATATATATATAT 4453
QY 1390 TTCGGCTGGCTTCAATAAATATATACAGACGAAATGAAATTTATATCATCTCACTTAA 1449
    || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 4454 ATATATGAAAAATATATATATATATATATATATGAAA-ATATATATATATATATATA 4512
QY 1450 ACGTAAATATGTTTAAAAAATCTGCTTTTATTCGATTCGATCAATGGGAAGAAATTT 1509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4513 TATATATATATATATATATATATATATATATGAAAAATATATATATATATATATG 4572
QY 1510 AATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1569
    || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 4573 AAATATATATATATATATATATATATATATATATATATATATATATATATATATA 4632
QY 1570 TATGAGAAATATACCATATGATATATATATATATATATATATATATATATATATAT 1629
    || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 4633 TATATATATATATATATATATATATATATATATATATATATATATATATATATA 4692
QY 1630 AAGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 1689
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4693 AATATATATATATATATATATATATATATATATATATATATATATATATATATATA 4752
QY 1690 ATGAAAAATTTATGATGAGAGGATTTGCAACTTAAGATATTTGAAA 1736
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4753 TATATATATATATATATATATATATATATATATATATATATATATATATATA 4799
```

RESULT 10

AAS46421/c ID AAS46421 standard; DNA; 13321 BP.

XX AAS46421;

DT 18-DEC-2001 (first entry)

DE Tumour suppressor gene derived chemically modified sequence #143.

XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;

KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;

XX cytosine methylation; ds.

XX Homo sapiens.

XX WO200168912-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-EP02955.

XX 15-MAR-2000; 2000DE-1013847.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPICENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-602752/68.

XX Claim 1; SEQ ID No 143; 27bp; English.

The invention relates to a nucleic acid comprising a sequence of 18 bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with bisulphite, of genes associated with tumour suppression and/or oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and 500 are missing from the sequence listing) sequences (Ss) and sequences complementary to (Ss). The nucleic acid may be a peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms and also to be used in an array for analysing diseases associated with CpG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases, by analysing cytosine methylations. The parameters may be compared to another set of genetic and/or epigenetic parameters, the differences serving as basis for diagnosis and/or prognosis events which are disadvantageous to patients. The present sequence is one of the 533 genomic sequences derived from tumour suppressor genes and oncogenes.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX Sequence 13321 BP; 2515 A; 580 C; 4121 G; 6105 T; 0 other;

Query Match 5.3%; Score 97.8; DB 22; Length 13321;

Best Local Similarity 47.6%; Pred. No. 2.2e-07;

Matches 288; Conservative 0; Mismatches 317; Indels 0; Gaps 0;

```
QY 1035 ATAGTATATAGATATATCCGTTAGATTATTAATATGTTAAACGAGTAAACATATA 1094
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2152 ATAAAGTTAAATATACGTTTAAATATACAACTTATCCCTCTTAATAATAATAA 2093
QY 1095 ACTATATATATTAATTCGAAATTTATTTGACAGTATATTTATATATATTAAGAGATAT 1154
    || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 2092 ATAAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2033
QY 1155 ATCTATTAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1214
    || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
DB 2032 ACAATATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1973
QY 1215 AGCAATTGAAAAATGAAATAGACACTTATTAATAAAAAAGCGGATGATTCAAATCTTCTTACA 1274
```


XX Claim 1: SEQ ID NO 240; 32pp + Sequence Listing; German.
 PS
 XX

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 CC
 XX

XX Sequence 6419 BP; 1565 A; 53 C; 1453 G; 3348 T; 0 other;

Query Match 5.2%; Score 96.2; DB 24; Length 6419;
 Best Local Similarity 48.8%; Pred. No. 3.9e-07;

Matches 352; Conservative 0; Mismatches 358; Indels 11; Gaps 3;

QY 912 ATACGAGACCAAAACAAAAATTAACMAAGAAAAATTTGAGCGCAAAATATTTT 971
 DB 2077 ATAAAGCAATTAATTAATTAATTAACAAAAATTAATAATTAACAAAAATTA 2018
 QY 972 GGAATTTTAAAGCGACTGCTGACGCACTTTGCCATATTAAACCTGACTATC 1031
 DB 2017 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1958
 QY 1032 TTATAGTTATAGATATATCCGTTAGATTAAAGTTAAAGTAAGAGTAATAACA 1091
 DB 1957 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1898
 QY 1092 ATAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1151
 DB 1897 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1842
 QY 1152 TATATCTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1211
 DB 1841 TCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1782
 QY 1212 AAAAGCATTAAGAAATGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1271
 DB 1781 AAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1722
 QY 1272 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1331
 DB 1721 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1662
 QY 1332 TATTAACCTCAATTTATCCATTGAGGAAGTTTGATCTTCTGATGTTGACACA 1391
 DB 1661 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1604
 QY 1392 CGGCTGGGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1451
 DB 1603 ----AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1548
 QY 1452 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1511
 DB 1547 AAACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1488
 QY 1512 TTACGATGAGAGTACATCAATCAAGAAATGATTAAGATTAAGATTAATTAACA 1571
 DB 1487 ATACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1429
 QY 1572 TGGAGAAATTAACCATATGATTAATTAATTAAGAAATATATGATTTGCAATTA 1631
 DB 1428 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1369
 QY 1632 G 1632
 DB 1368 G 1368

RESULT 15
 ABL70167/c

ID ABL70167 standard; DNA; 6154 BP.

XX ABL70167;

AC 01-JUL-2002 (first entry)

DE Chemically treated cell signalling DNA sequence#29.

KW Cell signalling; cytosine methylation; cell signalling disease;
 cancer; tumour; cytostatic; ds.

XX Unidentified.

PN WO200202807-A2.

PD 10-JAN-2002.

PF 29-JUN-2001; 2001WO-BE07471.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

PI WPI; 2002-154758/20.

PT Nucleic acid, useful for diagnosis and therapy of diseases associated
 with cell signalling e.g. cancer, comprises chemically modified genomic
 sequences of genes associated with cell signalling

PS Claim 1: SEQ ID NO 57; 24pp+sequence listing; English.

XX The invention relates to a nucleic acid comprising a sequence of at least
 CC 18 bases of a segment of chemically pretreated DNA of genes associated
 CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or PNA-oligomers for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.

XX Sequence 6154 BP; 1664 A; 146 C; 1358 G; 2986 T; 0 other;

Query Match 4.9%; Score 91.4; DB 24; Length 6154;
 Best Local Similarity 48.4%; Pred. No. 2.4e-06;

Matches 317; Conservative 0; Mismatches 331; Indels 7; Gaps 2;

QY 1064 TAAATATGTTAAAAACGATAAACAATTAATTAATTAATTAATTAATTAATTT 1123
 DB 689 TAAACAAAAACCTAAAAATTAATAAACAATCTACTATAATTAATAAATAAATTAAC 630
 QY 1124 GACAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1183
 DB 629 CAACAAATTAATAACCAACAATTAATAAACAATACTAAAAATTAATAAATAAATAA 570
 QY 1184 GAGTAATATATGATTTGTTATTAATAAAGCATTAAGAAATGAATTAAGAAATTA 1243
 DB 569 ATACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 510
 QY 1244 AAAAGACGCTGATCAATCTTTCTTCAATACTTGA----TTACTTTGATTAAGAT 1299
 DB 509 AAAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 450

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2003, 02:19:36 ; Search time 2425.86 Seconds
(without alignments)
11082.453 Million cell updates/sec

Title: US-08-876-132-1

Perfect score: 1660

Sequence: 1 AGACTACACAGCAAAAT.....AMGACTCCGACGACATCT 1660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_luv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
c 1	118	7.1	879 17	CNS01JRG
c 2	115.6	7.0	1101 17	CNS0021J
c 3	110.8	6.7	619 9	AL514935
c 4	110	6.6	1036 17	CNS031MJ
c 5	109.6	6.6	1025 17	CNS014J2
c 6	109	6.6	1036 17	CNS00599

7	109	6.6	1135 17	CNS033GQ	AL226115 Tetradon
c 8	106	6.4	924 13	BM415947	BM415947 OP21029 M
c 9	104.2	6.3	1223 17	B12981	B12981 F2AD11-Sp6
c 10	103.2	6.2	1885 10	BE420745	BE420745 HWM002.B0
c 11	103	6.2	964 13	BM416130	BM416130 OP21217 M
c 12	102.8	6.2	910 13	BM415636	BM415636 OP20714 M
c 13	102.2	6.2	907 17	CNS021J4	AL176953 Tetradon
c 14	101.2	6.1	974 17	CNS00ITT	AL075432 Tetradon
c 15	100.6	6.1	934 13	BM415088	BM415088 OP20159 M
c 16	100.4	6.0	948 13	BM415494	BM415494 OP20570 M
c 17	100.4	6.0	998 17	CNS00YXQ	AL096968 Drosophila
c 18	99.6	6.0	966 13	BM415686	BM415686 OP20766 M
c 19	99.4	6.0	982 13	BM415348	BM415348 OP20422 M
c 20	99.2	6.0	990 17	CNS006OI	AL065624 Drosophila
c 21	98.2	5.9	937 17	CNS006EL	AL062959 Drosophila
c 22	98.2	5.9	1101 17	CNS00EVL	AL069706 Drosophila
c 23	97.2	5.9	1139 17	AO897537	AO897537 HS_3153-A
c 24	96.6	5.8	660 17	BH183498	BH183498 023_L_07-
c 25	96.6	5.8	660 17	CNS070NJ	AL062049 T3 end of
c 26	96.6	5.8	804 17	AG077527	AG077527 Pan trogl
c 27	96.4	5.8	976 17	AG136173	AG136173 Pan trogl
c 28	95.8	5.8	1055 14	B0876453	B0876453 AGENCOURT
c 29	94.6	5.7	1101 17	CNS0039R	AL063932 Drosophila
c 30	94.2	5.7	838 17	AG061324	AG061324 Pan trogl
c 31	94.2	5.7	896 13	BM415641	BM415641 OP20719 M
c 32	93.6	5.6	811 17	CNS03YOO	AL266649 Tetradon
c 33	93.6	5.6	960 17	AG031602	AG031602 Pan trogl
c 34	93.6	5.6	1101 17	CNS00K85	AL077453 Drosophila
c 35	93.4	5.6	834 17	B12387	B12387 F21E20-Sp6
c 36	93.4	5.6	864 17	CNS04NBY	AL298519 Tetradon
c 37	93.2	5.6	1059 17	CNS002ZB	AL097133 Drosophila
c 38	93	5.6	1300 13	BM468018	BM468018 AGENCOURT
c 39	92.6	5.6	1185 12	BE273407	BE273407 CA_Eb001
c 40	92.4	5.6	915 12	BG520365	BG520365 PS20C09.Y
c 41	92.2	5.6	959 17	CNS00655	AL062806 Drosophila
c 42	91.8	5.5	1079 17	CNS035IX	AL228786 Tetradon
c 43	91.6	5.5	815 17	AG044049	AG044049 Pan trogl
c 44	91.6	5.5	1309 13	BM463041	BM463041 AGENCOURT
c 45	91.4	5.5	1021 17	AG032747	AG032747 Pan trogl

ALIGNMENTS

RESULT 1
CNS01JRG/ c 879 bp DNA linear GSS 12-JUN-2001
LOCUS
DEFINITION
Anopheles gambiae GSS T7 end of clone 14D07 of Notredame library from strain PEST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.
ACCESSION
AL147405
VERSION
AL147405.1 GI:7005551
KEYWORDS
GSS.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
REFERENCE
1 (bases 1 to 879)
Genoscope.
AUTHORS
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
JOURNAL
Submitted (16-FEB-2000) BBMT, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France
COMMENT
This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.
REFERENCE
2 (bases 1 to 879)
Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.
- Web : www.genoscope.cns.fr
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

[illegible][illegible]

FEATURES	Location/Qualifiers
source	1. .619

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Cl08B0122D08"
/clone_lib="ETL_NFL006_PL2"
/issue_type="Placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-Oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Peng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

BASE COUNT	52 a	8 c	10 g	514 t	35 others
ORIGIN					
Query Match		6.78;	Score 110.8;	DB 9;	Length 619

Best Local Similarity 46.8%; Pred. No. 3.2e-09;
Matches 280; Conservative 4; Mismatches 314; Indels 0; Gaps 0.

[illegible]

RESULT 4			
CNS03LMJ			
LOCUS	1036 bp	DNA	linear
DEFINITION	Tetradon nigriviridis genome survey sequence PUC-ori end of clone 03N11 of library G from Tetradon nigriviridis, genomic survey sequence.		
ACCESSION	AL250012	GI:7971024	
VERSION	AL250012.1		
KEYWORDS	GSS: genome survey sequence.		
SOURCE	Tetradon nigriviridis.		
ORGANISM	Tetradon nigriviridis		

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 1036)	Roest-Crollius, H., Jalllon, O., Dasilva, C., Bonneau, L., Fisher, C., Benoit, A., Fizames, C., Wincker, P., Brothier, P., Queller, F., Saurin, W. and Weissenbach, J.	Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1036)
ADDITIONAL	Reest-Crolius, H., Jallou, O., Dasilva, C., Fizes, C., Fisher, C.,
AUTHORS	Bonneau, V., Billaut, A., Quelet, F., Saurin, W., Betrol, A. and
TITLE	Weissenbach, J.
JOURNAL	Characterization and repeat analysis of the compact genome of the
ADDITIONAL	freshwater pufferfish <i>Tetraodon nigroviridis</i>
AUTHORS	Unpublished

REFERENCE 3 (bases 1 to 1036)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
 source
 1..1036
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_lib="G"
 /note="Genoscope sequence ID : C0B036CG06SP1-end ; PUC-Or1"

BASE COUNT 1005 a 6 c 7 g 17 t 1 others
 ORIGIN

Query Match 6.6%; Score 110; DB 17; Length 1036;
 Best Local Similarity 44.8%; Pred. No. 3.8e-09;
 Matches 422; Conservative 0; Mismatches 520; Indels 0; Gaps 0;

11 AAGCAAAATGAAAAATGATTAATTTCCAGGATTAAGCCGACTTAACAAAT 70
 38 AAGCAAAATGAAAAATGATTAATTTCCAGGATTAAGCCGACTTAACAAAT 97
 71 GAGTGAAG 130
 98 AAAAAAG 157
 131 AATTAAG 190
 158 AAAAAAG 217
 191 GCGATTTTAAATGGAATCGTGAAGAGAGAGAGAGAGAGAGAGAGAG 250
 218 AAAAAAG 277
 251 ATTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 310
 278 AAAAAAG 337
 311 TGACATTATGAAATACGTATATCAATTAATGGGGCTTCTATTTTAT 370
 338 AAAAAAG 397
 371 GATTGAAATCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 430
 398 AAAAAAG 457
 431 TTCAATGACTTTGGAATAATTCATGTGAGCGGAGGAACTTTGAAAT 490
 458 AAAAAAG 517
 491 TTGGAATTTGAAATATGAGCAAAAGAACTCAATGAGAGAGAGAGAG 550
 518 AAAAAAG 577
 551 AAGGAGATCGATATGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAG 610
 578 AAAAAAG 637
 611 GGAATAATCTATGCTGAGAGTGAATTAAGCAAAATTTGGAACACCGGT 670
 638 AAAAAAG 697
 671 TGTTTAGAAATGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730
 698 AAAAAAG 757
 731 AAGCAGCTCATATAGGAATTTATCAAAATTAAGTGAGCATAGTAATCGAG 790

Db 758 AAA 817
 Qy 791 TCTAGTACGAGACTATAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 850
 Db 818 AAA 877
 Qy 851 CTTTCTTTTGTATTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 910
 Db 878 AAA 937
 Qy 911 AAGAGTAAAGATTCGATACATGGAAGCAAAAGAGAGAGAGAGAGAGAG 952
 Db 938 AAA 979

RESULT 5
 CNS014J2/C
 LOCUS 1025 bp DNA linear GSS 26-JUL-1999
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
 BACN1111 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL104216
 AL104216.1 GI:5615827

KEYWORDS GSS.
 SOURCE Drosophila melanogaster.
 ORGANISM Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1025)
 Genoscope.

AUTHORS Direct Submission
 TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage ;
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr

COMMENT - Web : www.genoscope.cns.fr
 - Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.dbi.ac.uk> - . This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billard at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBeloBAC11.

FEATURES
 source
 1..1025
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="BACN1111"
 /clone_lib="DrosBAC"
 /plasmid="pBeloBAC11"
 /note="end : 77"

BASE COUNT 148 a 88 c 6 g 522 t 261 others
 ORIGIN

Query Match 6.6%; Score 109.6; DB 17; Length 1025;
 Best Local Similarity 35.0%; Pred. No. 4.4e-09;
 Matches 283; Conservative 98; Mismatches 427; Indels 0; Gaps 0;

69 ATGATGAAG 128
 890 KTTGGGRARADAAAAGGTGATGAGTAAATTTTATGATTAATGATTAARMA 831
 129 AAAAAAG 188
 830 AAAAAAG 771
 189 CTGATTTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 248
 770 KTTTGTGTGTAATAATATGTTGTAATTAATTAATTTTATGATTAARMA 711
 249 TAATTTGTTGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 308
 710 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 651

```

OY 309 TTGACATTATGTAATAGCTAGTATATCAATATAGGGGTTGCTATTATTTTGGC 368
Db 650 GAATWGTGTTGKTKKAAAAAAMAAAMAKWMTAAMKTAARAADGKTKWTKTTT 551
OY 369 AAGATTGAAATCTGAGTGAAGAAATAGTTTGCAGAGCAAAAAACCCCTGCCGTT 428
Db 590 RAAMAAKGRGCTGKAKGKTTRDWMWAAAAAAGGCGTGADARA 531
OY 429 TTTTCAATGACTTTGAAAAAATTCATTGTGAGCGTGACGAACCTTGAATTTTAA 488
Db 530 AKGDAAARADAGWMAAARAKRTAKTGKTKTKTGTACTAGAKTAAWTWMAAG 471
OY 489 CATGGAATTTGAAATTAAGCAAAAGCAAACTCAAAATATTTATTA 548
Db 470 TTGTTAKKMTATTTAAAAAAMAAAMAKWMTAAMKTAARAADGKTKWTKTTAA 411
OY 549 AAAAGAGATCGCATATGATTTTAAAGCAGAAACCTGACATGATGAAAAAAGAT 608
Db 410 AAAAATGTRWTRATKTKGTGDKAARAADAATTAATAAAAAAAGGCGTGADARA 351
OY 609 TTGGAATAATCTATCTGAGCTGATTAAGCAAAAAATTTGGAACTCAACCGGT 668
Db 350 AAAAAAANAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 291
OY 669 GTTGTGTAATGATGAGATGAAGAATGATGAAAAATTCACCTCGATTTAATGAA 728
Db 290 AAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 231
OY 729 GAAACAGCAGTCATATAGAAATTTTCAAAATTAAGTTGAGCATAGTAATGAGAT 788
Db 230 AAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 171
OY 789 GATCTGTAACGGCATATTAGAGTGGCAACTCGAAATGTTGATTAACAAGACT 848
Db 170 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 111
OY 849 GCCTTCGTTTGTATTTGTGAGAGAA 876
Db 110 CMAATACTTTTATTAATTAATTTNNNNNNNA 83

```

RESULT 6
CNS00599/c
LOCUS 1036 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence IETS end of BAC #
BACR1116 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL057797.1 GI:4932579
VERSION AL057797
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1036)
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT
determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> the BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Kammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's

PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 1036
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR1116"
/clone_lib="RPCI-98"
/note="end : TET3"

BASE COUNT 64 a 56 c 41 g 701 t 174 others

ORIGIN

Query Match 6.8%; Score 109; DB 17; Length 1036;
Best Local Similarity 41.0%; Pred. No. 5.5e-09;
Matches 317; Conservative 61; Mismatches 394; Indels 1; Gaps 1;
OY 1 AGATCTACACAGCAAAATGAAATAATTTTCGAGGTATTAAGCCGACT 60
Db 780 ACAATAAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
OY 61 TAAACAAATGACTGAAGAAAG-AAAAAATAATATATATTTTGAAGTAA 119
Db 720 AAAAAAGAAVMAAAAGSYHBAKTKCKCAAAAMATCMTAATAABKACTWAAAA 661
OY 120 GAGAAAGAAATTAAGAAAGACCTCGCTTAACAGTCAAAAACCAAAATTA 179
Db 660 AAATTCGAAAGAAAGAAAGATTAATATATATATATATATATATATATAT 601
OY 180 AGAAGAGACTGATTTTAAATGAAATGTAAGAAAGAAATTTTAAATTTT 239
Db 600 AAATTCGAAAGAAAGAAAGATTAATATATATATATATATATATATATAT 541
OY 240 TCAGGCAATTAATTTGTAAGTGAAGAAATTCAGTAAATAATTCAGATCA 239
Db 540 ATWAKRAAGAAANAAAGRTDGAKAAAAAATRTDAAAAAATTAATTAATTA 481
OY 300 TGTGTTGAATTTGACATTAATGAAATGATGATGATGATGATGATGATGAT 359
Db 480 GKEPTTBAAAAAABAGADAWMKADGGRGAGAKWCKRAGWTAATAATTAATTA 421
OY 360 TATTTGCGAGATTGAAATCTGAGTGAAGAAATAGTTTCGAGAGCAAAACC 419
Db 420 GATKRRARAGAKAAAAAANANAAAAAATTAATTAATTAATTAATTAATTA 361
OY 420 TTCCCGTTTTCATGACTTTGAAAAAATTCATTTGAGCGGTAGGAACTTGA 479
Db 360 AAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 301
OY 480 AATTTTTCATTTGGAATTTGAAAAATTAAGCAAAAGAACTCAAAATGGA 539
Db 300 AAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 241
OY 540 TATTTATTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 599
Db 240 AAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 181
OY 600 AAAAAATTTGGAATAATCTATGCTGAGAGTAAATTAAGCAAAATAATTT 659
Db 180 AAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 121
OY 660 CAACCCGGTGTTTTGAATGAGATGAAGAAATGTAAGAAATTAATCAACT 719
Db 120 AAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 61
OY 720 GTTAATGAAGAACAGAGCTCAATATAGAAATTTTCAAAATAAAGTTGA 772
Db 60 AAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 8
RESULT 7
CNS03360

LOCUS CENS03360 1135 bp DNA linear GSS 15-MAY-2000
 DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone
 208924 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL226115.1 GI:7885026
 VERSION AL226115.1
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorphi; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 REFERENCE 1 (bases 1 to 1135)
 AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Minckler,P., Brotlier,P., Queller,F.,
 Saurin,W. and Weissenbach,J.
 TITLE Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1135)
 AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1135)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 FEATURES
 SOURCE location/Qualifiers
 1..1135
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="208P24"
 /clone_1lb="G"
 /note="Genoscope sequence ID : C0AG208DH12SP1-end :
 PUC-ORI"

BASE COUNT 863 a 65 c 43 g 124 t 40 others

ORIGIN

Query Match 6.6%; Score 109; DB 17; Length 1135;
 Best Local Similarity 44.6%; Pred. No. 5.4e-09;
 Matches 358; Conservative 14; Mismatches 431; Indels 0; Gaps 0;

OY 7 ACACAAAGCAATTCGAAAAATAGATAAATTTTCGAGGATTTAAAGCCGACTTAAAC 66
 DB 216 AA 275
 OY 67 AAATGAGGAG 126
 DB 276 AA 335
 OY 127 AAAAAAAAAAG 186
 DB 336 AA 395
 OY 187 GACTGTGATTTTAAATGGAATTCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
 DB 396 AA 455
 OY 247 ATTAATTTGTTGTAAGTGAAGAAATCTAGATTAATAATGACATCAAAATGTGTTG 306
 DB 456 AA 515
 OY 307 AATTGACATTAATGTAATGATGATATATCAATTAATGGGGGTTTGTATTTTATTTG 366
 DB 516 AA 575

RESULT 8
 LOCUS BM415947/c 924 bp mRNA linear EST 28-JAN-2002
 DEFINITION OP21029 Mixed Stage EST's from Globodera pallida, the potato cyst
 nematode Globodera pallida CDNA, mRNA sequence.
 ACCESSION BM415947
 VERSION BM415947.1 GI:18382746
 KEYWORDS EST.
 SOURCE Globodera pallida.
 ORGANISM Globodera pallida
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 Tylenchoidea; Heterodidae; Heteroderinae; Globodera.
 REFERENCE 1 (bases 1 to 924)
 AUTHORS Heer,J., Sosinski,B., Pokrzywa,R.M., Warrly,A. and Opperman,C.
 TITLE Mixed Stage EST's from Globodera pallida, the potato cyst nematode
 JOURNAL Unpublished (2001)
 COMMENT Contact: Opperman, C
 Center for the Biology of Nematode Parasitism
 NC State University; IACR-Rothamsted
 Campus Box 7616; Raleigh, NC 27695, USA
 Tel: 919.515.6699
 Fax: 919.515.9500
 Email: warthog@unity.ncsu.edu
 GT11-8PCN_R_F10_GT11-8_R_079_ab1.
 FEATURES
 SOURCE location/Qualifiers
 1..924
 /organism="Globodera pallida"
 /db_xref="taxon:36090"
 /clone_1lb="Mixed Stage EST's from Globodera pallida, the
 potato cyst nematode"
 /note="Vector: lambda GT11; This is a collaborative effort
 between IACR-Rothamsted and North Carolina State
 University. The library was constructed from mixed stage
 G. pallida in lambda GT11 by Paul Burroughs,
 IACR-Rothamsted."

BASE COUNT 55 a 49 c 10 g 796 t 14 others

ORIGIN

Query Match 6.4%; Score 106; DB 13; Length 924;

OY 367 CGAATTTGAAATCTGAGTAAGAAATAGTTTGCAGAGACAAAAAACCCTTGCCGT 426
 DB 576 AA 635
 OY 427 TTTTTCGAATGACTTTGAGAAAAATTCATTTGTGAGCGGTAGCGAACTTTGAATTTT 486
 DB 636 AA 695
 OY 487 TACATTTGAAATTTTAAATAATTAAGCAAAAGAAATCTGAATGAAAAATATTATT 546
 DB 696 AA 755
 OY 547 AAAAAAGAGATTCGATTTGATTTTAAAGCAGAAAACTGACATTTGAATGAAAAAG 606
 DB 756 AAAAAAAAAAAAMTCYAAAAACAAATTAATAAAATATTAATTAATTAATTAATTA 815
 OY 607 ATTTCGAAAAATCTATCTGAGACTGAATTAAGCAAAAAATTTGGAACTCAACCG 666
 DB 816 AWACAAAAACACACACACAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 875
 OY 667 GTTGTGTTTGAATGACATGAAAGAAATGATGAAATATCAACTCGATGTTAATG 726
 DB 876 AAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 935
 OY 727 AAGAAACAGCAGGTCATATAGAAATTTTCAAAATAATTAAGTGAAGTAATCAG 786
 DB 936 AA 995
 OY 787 ATGATCTGATTAACGGGACTATTA 809
 DB 996 ACACATTAATAAAAAAAAAAAAAAAAAATATA 1018

Best Local Similarity 45.2%; Pred. No. 1.8e-08;
Matches 361; Conservative 0; Mismatches 438; Indels 0; Gaps 0;

```

OY 1 AGATCTACACAGCAATTTGAAAAATAGATTAATTTTGGCAGGTATTAAGCCGACT 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 807 AAAAAAAGAAAAAGAAAAATAAAAAAGAAAAAAGAAAAAAGAAAAAATAAA 748
OY 61 TAAACAAATGATGAGAGAGAAAAAATTAATATTTTGTAGTATTAAG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 747 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAATTTTAAAAA 688
OY 121 AGAAAAAATATATATATATATATATATATATATATATATATATATAT 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 687 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 628
OY 181 GAAAGAGCTGATTTTATGAAATCGAGAGAAAAAATTTTATTTTATTTT 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 627 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 568
OY 241 CGAGGATTAATTTGTTGATGATGAAAAATCTAGTAAAAATGAGATCAAAAT 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 567 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 508
OY 301 GTGTTGAAATTTGACATTTATGAAATACGATATATATATATATATATAT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 507 TAAAAAATATATATATATATATATATATATATATATATATATATAT 448
OY 361 ATTTTGCAGATTTGAAATCTGAGTGAAGAAATAGTTTGGCAGCAAAAAACC 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 447 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 388
OY 421 TGCGCTTTTTCATTAATGATTTGAAAAAATTCATTTGAGCGGTAGCAAACTTGA 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 387 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 328
OY 481 ATTTTACATTTGAAATTTGAAAAATTAAGCAAAAGAACTCAATGAAAAATAT 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 327 NNTTGGGGGGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 268
OY 541 ATTATTAAGAAAGAGATCGATATGATTTTAAAGCAAGAAAGTGCATTTGA 600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 267 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 208
OY 601 AAAAAATTTGAAAAAATCTATGCTGAGATGATTAAGCAAAAAATTTGGGAGCTC 660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 207 AAAAAAAGAAAAAAGAAAAAATTAAGCAAAAGAACTCAATGAAAAATAT 148
OY 661 AACCCTGCTTTTGAATGAGATGAAAGAAATGATGAAAAATATCACTCGATG 720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 147 ACTTAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 88
OY 721 TTAATGAAGAAAGAGATCAATATAGAAATTTTCAAAATTAAGTTGAGCATAG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 87 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 28
OY 781 AATCAGATGATCTACTAAC 799
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 27 AACAAAGNTGAGCTCGATC 9

```

RESULT 9
B12981/c
LOCUS T24D11-Sp6 Arabidopsis thaliana genomic clone T24D11, DNA
DEFINITION sequence.
ACCESSION B12981
VERSION B12981.1 GI:2094103
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

```

REFERENCE 1 (bases 1 to 1223)
AUTHORS Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
Ecker,J.
TITLE BAC End Sequences at ATGC
JOURNAL Unpublished (1997)
COMMENT Contact: Ecker J.
          Arabidopsis Thaliana Genome Center
          University of Pennsylvania
          Dept. of Biology, University of Pennsylvania, Philadelphia, PA
          19104
          Tel: 215-898-9384
          Fax: 215-898-8780
          Email: jecker@atgenome.bio.upenn.edu
          Seq primer: Sp6
          Class: BAC ends
          High quality sequence start: 199
          High quality sequence stop: 271.
          Location/Qualifiers
            source          1..1223
              /organism="Arabidopsis thaliana"
              /strain="Columbia"
              /db_xref="taxon:3702"
              /clone="T24D11"
              /clone_lib="TAMU"
              /sex="hermaphrodite"
              /note="Vector: BeLoBACII; Site_1: HindIII; Site_2: HindIII
              ; Produced by Rod Wing"
BASE COUNT 70 a 50 c 53 g 678 t 372 others
ORIGIN
Query Match 6.3%; Score 104.2; DB 17; Length 1223;
Best Local Similarity 30.5%; Pred. No. 3.4e-08;
Matches 295; Conservative 0; Mismatches 672; Indels 0; Gaps 0;
OY 7 ACACAAGCAATTTGAAAAATAGATTAATTTTGGCAGGTATTAAGCCGACTTAAC 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1220 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 1161
OY 67 AAATGATGAGAAAGAAAAAAGAAAAATTAATATATATTTGAGTTAGTAAAGAAAG 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1160 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 1101
OY 127 AAAAAATTAAGAAAGACTCGGCTTAACAGTCGAAAAACGAGAAATTAATTAAGAA 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1100 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 1041
OY 187 GACTGATTTTAAATGAAATCGTGAAGAAAGAAATTTTAAATTTTCAATTTTGA 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1040 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 981
OY 247 ATTAATTTGTTGATGATGAAAAATCTAGTAAATATGCAATCAAAATGCTGTG 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 980 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 921
OY 307 AATTGACATTTTGAATAGCTAGTATATCAATATGAGGGGTTGCTATTTTATTTG 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 920 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 861
OY 367 CGAAGATGAAATCTGAGTGAAGAAAGAAATTTGCGAGCAAAAAACCTTGGCGCT 426
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 860 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 801
OY 427 TTTTTCATATGACTTTGAAAAAATTCATTTGAGCGGTAGCGAACTTTGAATTTT 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 800 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 741
OY 487 TACATTTGAAATTTGAAAAATTAAGCAAAAGAACTCAATGAAAAATTTTATATA 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 740 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 681
OY 547 AAAAAAGAGATCGATATGATTTTAAAGCAAGAAAGTGCATTTGAATGAAGAAAG 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 680 AAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAAATTTTAAAAA 621

```

Oy		AATTTGGAAAATAAATCTATGCTGTAGACTGAATTAAAGAACAATAATCGCAACCCG	666
Dd		620 AAAAAAAAAANNNNNAANAIAAAAAAAAAAAAAIIINNAANNNNNNNAIINNAN	561
Oy		667 GTTGTGTTTTAGAAATGACGATGAAGAAATGATGAGAAAAATCAACTGCATGTTTAATG	726
Dd		560 ANAAA	501
Oy		727 AAGAACACGACGTCATATATAGGAATTTCTCAAATAATAAGTTGACATAGTAAATCAG	786
Dd		500 NAAAAAAAAANNNNNNNNANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	441
Oy		787 ATGATCTAGTAACGGACTATTAGAGTGTCGAACTCGAATAGTTTGATTAACAAGAA	846
Dd		440 NNN	381
Oy		847 GTCGCTTCGTTTTGTATTTGTGAGAGCATTCAGCACCTGAGAAAAAGAGCTGATATG	906
Dd		380 NNN	321
Oy		907 CACGACGAGTAAGACATTCCGATACAAITGAAAAGCAAAACCTAAAGAGCTTTTGAATTGA	966
Dd		320 NNNNNNNNNANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTTATCAGAAA	261
Oy		967 GTTTTGT 973 	
Dd		260 ATTITTAT 254	

BASE COUNT	1138 a	219 c	212 g	176 t	140 others
ORIGIN					

Query Match	6.28;	Score 103.2;	DB 10;	Length 1885;
Best Local Similarity	42.48;	Pred. No. 4.4e-08;		
Matches 384;	Conservative 0;	Mismatches 521;	Indels 0;	Gaps 0;

RESULT 11			
BM416130/c			
LOCUS			
BM416130	964 bp	mRNA	linear
			EST 28-JAN-2002

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Op21217 Mixed stage EST's from Globodera pallida, the potato cyst nematode Globodera pallida cDNA, mRNA sequence.	BM416130	BM416130.1	GI:18382929	EST.	Globodera pallida.
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Heteroderinae; Globodera.				
AUTHORS	1 (bases 1 to 964)				
TITLE	Heer, D., Sosinski, B., Pokrzywa, R. M., Wary, A. and Opperman, C.				
JOURNAL	Mixed stage EST's from Globodera pallida, the potato cyst nematode unpublished (2001)				
COMMENT	Contact: Opperman, C Center for the Biology of Nematode Parasitism NC State University; IACR-Rothamsted Campus Box 7616; Raleigh, NC 27695, USA Tel: 919.515.6699 Fax: 919.515.9500 Email: warthog@unity.ncsu.edu GILLPCNL_F06_1-30R_047.ab1.				
FEATURES	Location/Qualifiers				
source	1..964				
	/organism="Globodera pallida"				
	/db_xref="taxon:36090"				
	/clone_lib="Mixed Stage EST's from Globodera pallida, the potato cyst nematode"				
	/note="Vector: lambda GT11; This is a collaborative effort between IACR-Rothamsted and North Carolina State University. The library was constructed from mixed stage G. pallida in lambda GT11 by Paul Burroughs, IACR-Rothamsted."				
BASE COUNT	42 a	36 c	18 g	852 t	16 others
ORIGIN					
Query Match	6.2%	Score 103;	DB 13;	Length 964;	
Best Local Similarity	44.5%;	Pred. NO. 5.8e-08;			
Matches	424;	Conservative	0;	Mismatches 524;	Indels 5;
					Gaps 1;
123	AAAGCAAAAAAAGAGACCTCGGCTTAACAGTCGAAAAACCAAAATTAATAAAGA	182			
124		183			
125	AAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	904			
126	183 AAGAGACTGTGATTTTAATGAGAAATCGTGAGAAACAAATTTTAATTTTCG	242			
127		243			
128	903 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	844			
129	243 AGGATTAATTTGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT	302			
130		303			
131	843 AGAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	784			
132	303 GTTGATTTGACATTAATGAAATCGTGTATCAATTAATGGGGGTTTGTCTTTTAT	362			
133		363			
134	783 AAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	724			
135	363 TTTCGGAAGATGAAATCTGAGTGAAGAAATAGTTTGGAGAGCAAAAAACCTTG	422			
136		423			
137	723 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	664			
138	423 CCGTTTTTTCAATGACTTTGGAAAAAATTCATTGTGAGCGGTGCGAACTTTGAAT	482			
139		483			
140	663 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	609			
141	483 TTTTACATGCAATTTGAAAAAATAAGCAAAAGAACTCAATGAAATTAATTAAT	542			
142		543			
143	608 AAAAAAGAGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAA	549			
144	543 TATTAATAATAAGAGATCGATATGATTTTAAAAAGCAGAAACTGACATTGAATGA	602			
145		603			
146	548 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	489			
147	603 AAAGATTTGGAAAAATCTATGCTGAGAGTGAATTAAGCAAAAAAATTTGGAACTCA	662			
148		663			

[illegible]

[illegible]

AUTHORS	Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Winkler,P., Brotlier,P., Quetier,F., Sautin,W. and Weissenbach,J.									
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence									
JOURNAL	Unpublished									
REFERENCE	2 (bases 1 to 907)									
AUTHORS	Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Sautin,W., Bernot,A. and Weissenbach,J.									
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis									
JOURNAL	Unpublished									
AUTHORS	3 (bases 1 to 907)									
REFERENCE	Genoscope.									
AUTHORS	Direct Submission									
TITLE	Submitted (12-Apr-2000)									
JOURNAL	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .									
COMMENT	Location/Qualifiers									
FEATURES	1..907									
source	/organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="225M09" /clone_11b="G" /note="Genoscope sequence ID : C0AG225CG05SP1-end : PUC-Orl"									
BASE COUNT	27 a 30 c 22 g 759 t 69 others									
ORIGIN										
Query Match	6.2%; Score 102.2; DB 17; Length 907;									
Best Local Similarity	43.0%; Pred. No.86-08;									
Matches	332; Conservative 42; Mismatches 393; Indels 5; Gaps 1.									
QY	1	AGATCTACACAAGCGCAATTTGAAAAATGATGATAAATTTTCCAGAGTATTAAAGCCGACT	60							
DB	781	AAAKADAAAAAAMAAAAACWAAAAAAMAAAAAMWSASAAAAAAMAAAAAAGCTRACT	722							
QY	61	TAAACAATGAGTG-----AAGAAGAAAAAGAAAAATTAATACATATTTGAGTTAGT	115							
DB	721	AAAAAAMAGCGGBGTGRTAKRAAARRGAAAAAAMMBWATCACAAMATATAA	662							
QY	116	AAAAAGAGAAGAAAAATTAAGAAGACCTCGCTTAACAGTCGAAAAACGAAATAAT	175							
DB	661	AMABRAGAAAAAAMAAAAAMAAAVAAAAACAABAAAAACACSGAKAAAAAGAAAAAA	602							
QY	176	AAAAAGAAAGACGCTGATTTTAAATGGAATCGTGAGAAAAAGAAATTTTAATTTTC	235							
DB	601	AAAAAAATCAGHAAAAAASMAAAATTAAMAAAAABGAAAAAAMAAAAAAMAAAAA	542							
QY	236	ATTTCGAGGGCTTAATTTGTTGTAAGTTGATGAAAAATCTGATTAAMAAATGACATCA	295							
DB	541	AAACAAAAABGAAAGAAABAAAATATGWMAMSMACYNASAAAMSCAAAAABACATRA	482							
QY	296	AAATGCTGTGAATTTGACATATTGCAATACGTATATCAATTAATGGGGTTGTCT	355							
DB	481	AAAAAAMAAAAAMKRAAAAAAAMAAATHTYACAAAAATACYGACAAATAAAAA	422							
QY	356	ATTATTATTTGGGAAGATTGAANAATCGAGTGAAGAAATATGTTGGCAGACGCAAAAA	415							
DB	421	TAAAAAAMAAAAAMAAKGAAMAAAAAMAAARGAAMAGGAGNADDDGAAAAAA	362							
QY	416	ACCCTGGCGTTTTCCTCAATGACTTTGGAAAAAATTCATTTGTGACGGGTGCGAAACT	475							
DB	361	AAATCTATYTGABAHCAATDAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA	302							
QY	476	TTGAATTTTTCACATTGGAATTTGAAAAAATTAAGCAAAAAAGAACTCAATTGAAAAA	535							
DB	301	AAA	242							
QY	536	ATATTATTATAAAAAAGAGATCGATGATGATTTTAAAGCAGAAAACTGACATTGAA	595							

FEATURES
sourceLocation/Qualifiers
1. .934

/organism="Globodera pallida"
/db_xref="taxon:36090"
/clone_lib="Mixed Stage EST's from Globodera pallida, the
potato cyst nematode"
/note="Vector: Lambda GT11; This is a collaborative effort
between IACR-Rothamsted and North Carolina State
University. The library was constructed from mixed stage
G. pallida in Lambda GT11 by Paul Burroughs,
IACR-Rothamsted."

BASE COUNT 45 a 80 c 17 g 777 t 15 others
ORIGIN

Query Match 6.1%; Score 100.6; DB 13; Length 934;

Best Local Similarity 45.6%; Pred. No. 1.5e-07;
Matches 360; Conservative 0; Mismatches 429; Indels 1; Gaps 1;

```
QY 9 ACAAGCCAAATTGAAAAATAGATAAATTTTCGACGATTAAGCCGACTTAACAA 68
Db 792 ATAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 733
QY 69 ATGAGTGAAGAGAAAGAAAAAATTAATACATATTTTGAAGTAAAGAAAGAA 128
Db 732 AAGAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 673
QY 129 AAAATTAAGAAGACCTCGCTTAACAGTCAAAAAACGAATATATAAAGAAAGA 188
Db 672 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 613
QY 189 CTGCAATTTTAATGAAATCGTGAAGAAAGAAATTTTAATTTTCATTTTCGAGGAT 248
Db 612 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 553
QY 249 TAAATTTGTTGATGAAATCTAGATAAATAATGCATCAAAATGCTGTGA 308
Db 552 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 493
QY 309 TTTGACATTAATTAAGATAGTATATCAATATGAGGCTTGTCTATTTTATTTGCG 368
Db 492 TAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 433
QY 369 AAGATTGAATCTGAGTGAAGAAATATGTTGGAGAGCAAAAAACCTTGCCTTT 428
Db 432 AAAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA- AATTA 374
QY 429 TTTCAATGACTTTGAAAAAATTCATTGTGAGCGGTAGCGAACTTTGAATTTTTTA 488
Db 373 GAAAAAAGAAAAAGTGAATGTGCTTAATTAACNCCNANANNNNNNTTGGGGGGG 314
QY 489 CATTGGAATTTGAAAAATTAAGCAAAAGAACTCAATGGAATAAATATTTATTTAA 548
Db 313 GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 254
QY 549 AAAAGCAGATCGGATATGATTTTAAAGCAGAAATGACATTGAATGAAAAAAGAT 608
Db 253 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 194
QY 609 TTGAAAAAATCTATGAGAGTGAATTAAGCAAAAAAATGGAACCTCAACCGGT 668
Db 193 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 134
QY 669 GTTGTTTTGAAGATGACATGAAGAAATGATGAATAATATCACTCGATGTTAATGAA 728
Db 133 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 74
QY 729 GAAAGCAGAGTCAATATAGCAATTTCAAAAAATTAAGTTGACATAGTAATCAGAT 788
Db 73 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 14
QY 789 GATCTAGTAA 798
Db 13 TATCNACTRA 4
```

Search completed: July 17, 2003, 06:38:51
Job time : 244.36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2003, 02:19:35 ; Search time 79.9943 Seconds

(Without alignments)
6363.996 Million cell updates/sec

Title: US-08-876-132-1

Perfect score: 1660
Sequence: 1 AGATCTACACAGCAAAATT.....AAGACTCCGACGACATCT 1660Scoring table: IDENTITY_NNC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCBUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	81.2	4.9	7218	1	US-08-232-463-14
C 2	69.6	4.2	19124	2	US-08-487-826B-13
C 3	67.8	4.1	5852	1	US-07-867-106-2
C 4	62.6	3.8	615	4	US-08-998-416-186
C 5	60.2	3.6	1956	4	US-08-559-896B-1
C 6	59.9	3.6	6243	2	US-09-056-075-1
C 7	57.6	3.5	837	4	US-08-998-416-288
C 8	57.6	3.4	240	1	US-08-628-417-6
C 9	56.6	3.4	2570	2	US-09-056-075-2
C 10	55.8	3.4	636	4	US-08-998-416-1137
C 11	55.4	3.3	356	2	US-08-520-678A-22
C 12	55.4	3.3	356	4	US-08-897-126-22
C 13	55.4	3.3	9646	3	US-08-811-566-1
C 14	55.4	3.3	9646	4	US-09-034-756-1
C 15	54	3.3	658	4	US-08-998-416-595
C 16	54	3.3	2447	2	US-09-014-969-14
C 17	53.8	3.2	8920	2	US-08-446-855A-1
C 18	53.8	3.2	8920	4	US-09-150-741-1
C 19	53	3.2	4766	5	PCT-US93-07261-10
C 20	52.8	3.2	3275	4	US-09-370-838-151
C 21	52.6	3.1	6152	4	US-08-973-462-1
C 22	51.6	3.1	12980	3	US-08-811-566-5
C 23	51.6	3.1	12980	4	US-08-034-756-5
C 24	51	3.1	1798	4	US-09-797-906-1
C 25	50.8	3.1	2030	4	US-09-512-342-1
C 26	50.6	3.0	72604	4	US-09-268-992-7
C 27	50.6	3.0	72604	4	US-09-657-474-7

28	50.2	3.0	2674	4	US-09-817-180-1	Sequence 1, Appl1
29	49.4	3.0	2223	1	US-08-257-073-4	Sequence 4, Appl1
C 30	49.2	3.0	3718	4	US-09-424-283-6	Sequence 6, Appl1
C 31	48.8	2.9	6265	4	US-09-129-112-3	Sequence 3, Appl1
C 32	48.8	2.9	13737	4	US-09-538-414-10	Sequence 10, Appl1
C 33	48.6	2.9	20674	4	US-09-641-638-651	Sequence 1, Appl1
C 34	48.4	2.9	51952	3	US-08-947-823-1	Sequence 24, Appl1
C 35	48	2.9	1493	1	US-08-340-820-24	Sequence 2, Appl1
C 36	48	2.9	1493	1	US-08-593-535-24	Sequence 33, Appl1
C 37	47.8	2.9	1117	4	US-09-247-373B-33	Sequence 13, Appl1
C 38	47.8	2.9	5852	1	US-07-867-106-2	Sequence 9, Appl1
C 39	47.8	2.9	19124	2	US-08-487-826B-13	Sequence 1, Appl1
C 40	47.6	2.9	3000	1	US-08-764-100-9	Sequence 7, Appl1
C 41	47.6	2.9	168575	4	US-09-426-290-1	Sequence 4, Appl1
C 42	47.4	2.9	394	2	US-08-623-906A-7	Sequence 27, Appl1
C 43	47.4	2.9	1066	1	US-08-157-101A-4	
C 44	47.4	2.9	3188	1	US-08-017-664-4	
C 45	47.2	2.8	1447	4	US-09-443-041A-27	

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F1s
US-08-232-463-14
Query Match 4.9%: Score 81.2: DB 1: Length 7218:


```

Db      428 TCAACATATATTTATTAATAAGATATTTATATAATAATAATTTTCAATATTTTAA 369
Oy      241 CGAGGATTAATTTGCTGTAGTGAATAATCTAGATTAATAATGAGATCAAAAT 300
Db      368 TAAATTAATCTTTATTAATAATAATATATTTATTAATAACAATTAATAATAT 309
Oy      301 GTGTGAATTTGACAT-TATTGAATACGTACTATATCAATAATAGGGGTTGTCTATT 359
Db      308 ATTAATATTTGATATCTATTAATAATTAATTAATAAGAAATATATATCTAATAATAT 249
Oy      360 TTTTTCGCAATGAAATCTGAGTGAAGAAATAGTTGCGAGACCAAAAAACCC 419
Db      248 TTTAACTAATTTTAAATTTGAACATGACTAAATAGTATTCATTTAAATATTTT 189
Oy      420 TTGCGCTTTTTCCTCAATGACTTTGAAAAAATTCATTTGACCGGTAGCAACTTGA 479
Db      188 TTATATATATATAAATATTAATAATGATGATTAATTAATTAATAATATATAT 129
Oy      480 AATTTTTCATTTGAAATTTGAAAAATTAAGCAAAAGAACTCAATGAAAAATAT 539
Db      128 AAGTATTTATTAATCAATATTAATTTATTAATAATGATTAATTAATTAATTAAT 69
Oy      540 TTTTATTAATAAAGAGATCGATATGATTTTAAAGCAGAAACTGCAATGAAT 596
Db      68 AATACTTAATATTTTATTAATAAAGCTTATATTAATCTTTTAAATTAATTAAT 12

```

RESULT 5

```

US-08-559-896B-1
Sequence 1, Application US/08559896B
Patent No. 6310046

```

```

GENERAL INFORMATION:
APPLICANT: Patrick E. Duffy
APPLICANT: Christian F. Ockenhouse
TITLE OF INVENTION: SEQUESTRIIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: John Moran
STREET: USA MRMC - MCMR-JA
CITY: FORT DETRICK, FREDERICK
STATE: MARYLAND
COUNTRY: USA
ZIP: 21702-5012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/559, 896B
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Moran, John
REGISTRATION NUMBER: 26,313
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-7714
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1956 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Double
TOPOLOGY: Linear
US-08-559-896B-1

```

```

Query Match      3 6%; Score 60.2; DB 4; Length 1956;
Best Local Similarity 41.9%; Pred. No. 0.00032;

```

	Matches	Conservative	0;	Mismatches	598;	Indels	6;	Gaps	1;
Oy	25	AAATAGATAAATTTTCCGAGTATTAAGCCGACTTAAGCAATGAGTGAAGAGAA							84
Db	884	ATAGAGAAAAATTTACCAAAATTTGAAAAAGAGAAAAATTAATAATGATTAAGATCAAA							943
Oy	85	ACAAAAAATTAATACATATTTTGGACTTGGTAAAGAGAGAAAAAATTAAGAGAACCC							144
Db	944	TGATATTAATTAATATGAGAGAAATTAACAAAATGATGTGATGAAATTCACATGTAA							1003
Oy	145	TCGGCTTACAGTCGAAAAACCGAATATTAAGAAAGAGACGTGATTTTAAATGG							204
Db	1004	GGAGACCAATTTACAGATATACAAAAAGAAAAATACAAAATTTAGAACTAGAGAAA							1063
Oy	205	AAATCGTGAGAAAAAATTTTAAATTTTCATTTTTCGAGGATTAATTTGTGTAAGTT							264
Db	1064	TCGATAGACTTTATTAAGCAATTTAGATAGATGATGAGAACGACGATGAATCC							1123
Oy	265	GATGAAAAATCTAGATTAATAATGCAATCAAAATGTCTGAATTTGCAATTTGAAA							324
Db	1124	CCATGAGAAATTTAAGCAGAAATGAAAAAGATTAATTTACATGAAATTTAAAAATG							1183
Oy	325	TACGATATATCAATTAATGGGGTTTGTCTATTTTATTTTTCGAAATTCGAA							384
Db	1184	AATCTACCAAAAAATTAAGAGAAAAATGTAATGATTTTATTAATACGACAAATATG							1243
Oy	385	GTGAAAAAATAGTTTTCGAGAGCAAAAAACCCCTGCGGTTTTCGAAATGACTTG							444
Db	1244	ATAGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA							1303
Oy	445	GAAAAAATTCATTTGAGCGGTAGCAAACTTTGAATTTTTCATTTGCAATTTGAAA							504
Db	1304	CAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA							1363
Oy	505	AAATAGCAAAAGAACTCAATGCAAAAAATTAATTAATTAATTAATTAATTAATTA							564
Db	1364	ATTAAGAGCCCAAGATAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTA							1423
Oy	565	TGATTTTAAAGCAGAAACTGCACTTAATGAAGAAAAAAGATTTGAAAAAATCTATG							624
Db	1424	AAAAATGTAATTTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT							1483
Oy	625	CTGAGAGTAATTAAGCAAAAAAATTTGGAACTCAACCGGTGTTTGAAGAATGA							684
Db	1484	CACCAATTCACATTAAGAGAAATCAAGTATGATGTTCTCAGAAAAATTAATTCAGATTA							1543
Oy	685	CGATGAAGAAATGATGAATAATATCAACCTGATGTTTAATGAAGAACGACGATCAAT							744
Db	1544	TTCAAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA							1603
Oy	745	ATAGGAATTTATCAAAAAATTAAGTTGACATATTAATCAATGATCTAGTACAGGAGC							804
Db	1604	ATTAAGAAATTAACAAATATTTGATTAATAAAAAATTTGATGATTAATAAAAAATTTG							1663
Oy	805	TATTAGAGTGTGAATCGAATAGTTTGAATAAACAAGAGTGCCTTTCTTTTGA							864
Db	1664	TAAAAAGTCTTGAGATTAATAAAGTGTGATGATTAATAACAATGTTGATGATTAATAA							1723
Oy	865	TTTGAGAGAAATTCAGCACT-----GAGAAAGAGCTGATTAATGCAAGAGATA							918
Db	1724	ATCTTGATCGTATTAATAAATGTTGATGATTAATAAATGTTGATGATTAATAAATGTTG							1783
Oy	919	AAATTTGATCAATTAAGAGCAAAACTAAAGAGCTTTGAATTTGAGTTTGT							978
Db	1784	GAGATATTAACAATCTGAGATTAACAATAATGCTGAGATTAATAACAATGCTGAGATA							1843
Oy	979	ATAAGATTTTGTGAGTGAATAATAGCAATTAATGATGATATTTCTCAACCAAAAAAG							1038
Db	1844	TAAACAAATCTGTACATTAATTAACAGTTGAACATATAGACGAGCGGAAAAAACCA							1903
Oy	1039	ACTGCAAGTAAAGAA 1057							
Db	1904	ATCTTGATTAATCCAAAAA 1922							

RESULT 6

US-09-056-075-1
Sequence 1, Application US/09056075
Patent No. 5953368
GENERAL INFORMATION:
APPLICANT: Johnson, Eric A.
APPLICANT: Bradshaw, Marite
APPLICANT: Rood, Julian
TITLE OF INVENTION: Expression System for Clostridium
NUMBER OF SEQUENCES: 2
TITLE OF INVENTION: Species
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,075
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95238
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc.feature
LOCATION: 3770..4013
OTHER INFORMATION: /note="RP4 origin of DNA transfer (orit) from
OTHER INFORMATION: plasmid RP4"
US-09-056-075-1

Query Match 3.6%; Score 59; DB 2; Length 6243;

Best Local Similarity 51.7%; Pred. No. 0.00076;
Matches 134; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

9 ACAGGCAATGTAATAATTTTCGCGAGTATTAAGCCGCTTAACAA 68
1192 AA 1251
69 ATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 128
1252 AA 1311
129 AAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 188
1312 AA 1371
189 CTGATTTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 248
1372 ATATATAATAATAATAATAATAATAATAATAATAATAATA 1431
249 TAATTTGTTGAGTTGAT 267
1432 TTTTATATATATATAT 1450

RESULT 7

US-08-998-416-288/C
Sequence 288, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 288:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: PAG1241RP
ORGANISM: PAG1241RP
US-08-998-416-288

Query Match 3.5%; Score 57.6; DB 4; Length 837;

Best Local Similarity 44.8%; Pred. No. 0.00093;
Matches 260; Conservative 0; Mismatches 319; Indels 1; Gaps 1;

18 ATTGAAAAATAGATAAAATTTTCGCGAGTATTAAGCCGAGTAAACAAATGAGTGA 77
591 AATTAAATATAGATAAAATTTTAATTAATTAATTAATTAATTAATTAATTAATTA 532
78 GAAG 137
531 AAGTGAAT 472
138 GAAG 197
471 TAATTTTAAATTAATTTCTTAATAAGAGTTAAATTAATTAATTAATTAATTAAT 412
198 TTAATGAG 257
411 AAGATGAT 352

OY 258 GTAAGTGTGATAAAATCTAGATATAAAATGCAGATCAAAAATGTGTAATTGACAT- 316
DB 351 ATAAATAAATAATATATATTTTATATAACAAATTAATAATATATATATATATATATC 292
OY 317 TATTTGAATAGTACTATATATCAATATATGGGGTTTGCTATTTTATTTTGGCAAGATTGA 376
DB 291 TATTTAAATAATTTTATATAAGAAATAATAATATCTAAATATTTTAACTAATTTAA 232
OY 377 AAATCTGAGTGAAGAAATAATAGTTGCGAGAGCAAAAACCTGCGCTTTTTCGAA 436
DB 231 AATTGACATGACTATATATAGTATATTCATATTTAAATATATTTTATATATATAATA 172
OY 437 TGACTTTGCAAAAATTCATTTGAGCGGTAGCGCAACTTTGAAATTTTTCACATTGAA 496
DB 171 TTAATTAATGATGAATTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTA 112
OY 497 ATTGCAAAAATTAAGCGCAAAAACCTCAATGCAAAAATTTTATATAAAAAGAG 556
DB 111 TAAATTAATTTTATTAATTAATGATTAATTAATTAATTAATTAATTAATTAATTA 52
OY 557 ATCGATATGATTTTAAAGCAGAAACTGACATTGAAT 596
DB 51 ATAAAAAGTTTATTTAATCTTTTAAATTAATTAATTAAT 12

RESULT 8

US-08-628-417-6

Sequence 6, Application US/08628417

Patent No. 5627054

GENERAL INFORMATION:

APPLICANT: GILLESPIE, DAVID

TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC

TITLE OF INVENTION: POLYMERASE CHAIN REACTION

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL

ADDRESS: DEFENSE COMMAND

STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)

CITY: ABERDEEN PROVING GROUND

STATE: MARYLAND

COUNTRY: USA

ZIP: 21010-5423

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/628.417

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BIFFONT, ULYSSES J

REGISTRATION NUMBER: 39,908

REFERENCE/DOCKET NUMBER: DAM 398-94

TELECOMMUNICATION INFORMATION:

TELEPHONE: 410-671-1158

TELEFAX: 410-671-2534

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 240 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: oligodeoxynucleotide

HYPOTHETICAL: NO

ANTI-SENSE: YES

US-08-628-417-6

Query Match

Best Local Similarity 54.5%; Score 57; DB 1; Length 240;

Matches 114; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

OY 16 AAATGAAAAATAGATATAAAATTTTCGAGGTATTAAGCCGACTTAACAAATGACTG 75
DB 12 ACAATTAACCTTGAATAATTTTACTAAAAAATAAATAAATAAATAAATAAATAA 71
OY 76 AAGAAGAAAAAATAAATAATATATTTTGTAGTGTAAAGAGAAAAAATAA 135
DB 72 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 131
OY 136 AAGAAGCCTGCGCTTACAGTCGAAAAACCAATAATTAATAAAGAAGAGACTGTGAT 195
DB 132 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 191
OY 196 TTTTAATGAAATCGTGAGAAAGCAAA 224
DB 192 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 220

RESULT 9

US-09-056-075-2

Sequence 2, Application US/09056075

Patent No. 595368

GENERAL INFORMATION:

APPLICANT: JOHNSON, Eric A.

APPLICANT: Bradshaw, Marile

TITLE OF INVENTION: Expression System for Clostridium

TITLE OF INVENTION: Species

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/056.075

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J

REGISTRATION NUMBER: 27386

REFERENCE/DOCKET NUMBER: 960296.95238

TELECOMMUNICATION INFORMATION:

TELEPHONE: 608-251-5000

TELEFAX: 608-251-9166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2570 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-09-056-075-2

Query Match 3.4%; Score 56.6; DB 2; Length 2570;

Best Local Similarity 44.8%; Score 56.6; DB 2; Length 2570;

Matches 218; Conservative 0; Mismatches 269; Indels 0; Gaps 0;

OY 170 AATAATAAAGAAAGAGAGCTGATTTTAAATGAATCGTGAAGAAAGAAATTTTA 229
DB 116 AACATATACATATCAATATATTTTGTAAACCTAAATTAATATATCAAAATTTT 175
OY 230 AATTTCATTTTCGAGGATTAATTTGCTAGTGAATGAAATCTAGATTAATAAATGC 289
DB 176 ATTAGATGTTTACATATATGATTAATTTGCTAAATAAATGCGCTTATTAATAATTTA 235

1
2 ZIP: 07802
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: 5.25 disk
5 OPERATING SYSTEM: IBM PC compatible
6 SOFTWARE: Patent in Release #1.0, Version #1.30
7 CURRENT APPLICATION DATA:
8 APPLICATION NUMBER: US/08/811,566
9 FILING DATE: 03-MAR-1997
10 CLASSIFICATION: 435
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Jackson Esq., David A.
13 REGISTRATION NUMBER: 26,742
14 REFERENCE/DOCKET NUMBER: 1113-1-006
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: 201-487-5800
17 TELEFAX: 201-343-1684
18 INFORMATION FOR SEQ ID NO: 1:
19 SEQUENCE CHARACTERISTICS:

LENGTH: 9646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-811-566-1

Query Match 3.3%; Score 55.4; DB 3; Length 9646;
Best Local Similarity 56.2%; Pred. No. 0.0049;
Matches 104; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 10 CAAGGCAATTTGAAAAATAGATAAATTTTCGAGGTATTAAGCCGACTTAACAA 69
DB 9585 CACAGCTAGCCGCTAGCTAGGCTAAGATGAGCCACCATTAAGAAAGAAAGAA 9526
QY 70 TGAAGTGAAG 129
DB 9525 GGAAG 9466
QY 130 AATATAAG 189
DB 9465 AAAGAGAGAGAG 9406
QY 190 TGTGA 194
DB 9405 TAAGA 9401

RESULT 14
US-09-034-756-1/c
Sequence 1, Application US/09034756
Patent No. 6392028

GENERAL INFORMATION:
APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/034,756
FILING DATE: 04-May-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-09-034-756-1
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Query Match 3.3%; Score 55.4; DB 4; Length 9646;
Best Local Similarity 56.2%; Pred. No. 0.0049;
Matches 104; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 10 CAAGGCAATTTGAAAAATAGATAAATTTTCGAGGTATTAAGCCGACTTAACAA 69
DB 9585 CACAGCTAGCCGCTAGCTAGGCTAAGATGAGCCACCATTAAGAAAGAAAGAA 9526
QY 70 TGAAGTGAAG 129
DB 9525 GGAAG 9466
QY 130 AATATAAG 189
DB 9465 AAAGAGAGAGAG 9406
QY 190 TGTGA 194
DB 9405 TAAGA 9401

RESULT 15
US-08-998-416-595/c
Sequence 595, Application US/08998416
Patent No. 6239264

GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPIT
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264ch Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996

ATTORNEY/AGENT INFORMATION:
NAME: Weigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/GC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 595:
SEQUENCE CHARACTERISTICS:
LENGTH: 658 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1408RP

US-08-998-416-595

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2003, 02:22:20 ; Search time 612.027 Seconds
(without alignments)
5635.000 Million cell updates/sec

Title: us-08-876-132-1

Perfect score: 1660
Sequence: 1 AGATCTACACAGCAAAATT.....AAGGACTCCGACGACATCT 1660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 144851 seqs, 1038787357 residues

Total number of hits satisfying chosen parameters: 2897702

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/PCTUS07_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq1.*
11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq3.*
13: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.8	5.0	9539	US-10-239-676-52	Sequence 52, App1
2	83.6	5.0	7823	US-10-239-676-198	Sequence 198, App1
3	74.2	4.5	18997	US-10-172-086-17	Sequence 17, App1
4	71	4.3	9539	US-10-239-676-51	Sequence 51, App1
5	70	4.2	411	US-09-960-352-14521	Sequence 14521, A
6	69.4	4.2	446	US-09-960-352-3400	Sequence 3400, Ap
7	69	4.2	821	US-09-878-574-4653	Sequence 4653, Ap
8	69	4.2	960	US-10-198-846-6381	Sequence 6381, Ap
9	67.4	4.1	327	US-09-960-352-4630	Sequence 4630, Ap
10	67.4	4.1	7657	US-10-239-676-185	Sequence 185, App
11	66.6	4.0	428	US-09-960-352-573	Sequence 573, App
12	65.4	3.9	12405	US-10-239-676-35	Sequence 35, App1
13	64.8	3.9	4985	US-10-094-240-10	Sequence 10, App1
14	64.2	3.9	5689	US-10-239-676-90	Sequence 90, App1
15	64.2	3.9	6098	US-10-239-676-126	Sequence 126, App
16	63.4	3.8	529	US-09-983-965-2109	Sequence 2109, App

C 17	63.4	3.8	6665	14	US-10-239-676-4	Sequence 4, App1
C 18	63.2	3.8	525	14	US-10-198-846-1483	Sequence 1483, Ap
C 19	62.8	3.8	6352	14	US-10-172-086-24	Sequence 24, App1
C 20	62.8	3.8	1223197	14	US-10-027-632-179264	Sequence 179264, A
C 21	62.4	3.8	6306	14	US-10-239-676-224	Sequence 224, App
C 22	62.2	3.7	5314	14	US-10-155-533-1	Sequence 1, App1
C 23	62	3.7	15832	14	US-10-239-676-106	Sequence 106, App
C 24	61.4	3.7	408	11	US-09-960-352-1221	Sequence 1221, Ap
C 25	61.2	3.7	516	11	US-09-960-352-5785	Sequence 5785, Ap
C 26	61	3.7	479	11	US-09-960-352-12872	Sequence 12872, A
C 27	60.8	3.7	344	11	US-09-960-352-1036	Sequence 1036, Ap
C 28	60.8	3.7	17848	14	US-10-239-676-28	Sequence 28, App1
C 29	60.6	3.7	424	11	US-09-960-352-11218	Sequence 11218, A
C 30	60.6	3.7	5127	14	US-10-239-676-132	Sequence 132, App
C 31	60.4	3.6	390	11	US-09-960-352-3640	Sequence 3640, Ap
C 32	60.4	3.6	431	11	US-09-960-352-5558	Sequence 5558, Ap
C 33	60.4	3.6	4237	10	US-09-745-763-20	Sequence 20, App1
C 34	60.2	3.6	1267	14	US-10-001-843-45	Sequence 45, App1
C 35	60.2	3.6	1267	15	US-10-001-843-45	Sequence 45, App1
C 36	60.2	3.6	1956	10	US-09-351-794-1	Sequence 1, App1
C 37	60.2	3.6	7657	14	US-10-239-676-185	Sequence 185, App
C 38	59.8	3.6	2000	11	US-09-938-842-3403	Sequence 3403, Ap
C 39	59.8	3.6	15732	14	US-10-239-676-96	Sequence 96, App1
C 40	59.6	3.6	451	11	US-09-960-352-10262	Sequence 10262, A
C 41	59.6	3.6	640681	11	US-09-790-988-1	Sequence 1, App1
C 42	59.4	3.6	6815	14	US-10-239-676-50	Sequence 50, App1
C 43	59.4	3.6	11047	14	US-10-239-676-187	Sequence 187, App
C 44	59.4	3.6	12465	14	US-10-239-676-31	Sequence 31, App1
C 45	59.2	3.6	283	11	US-09-960-352-9095	Sequence 9095, Ap

ALIGNMENTS

RESULT 1
US-10-239-676-52/c
: Sequence 52, Application US/10239676
: Publication No. US20030082609A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
: FILE REFERENCE: 5013.1003
: CURRENT APPLICATION NUMBER: US/10/239, 676
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: PCT/EP01/03968
: DE 10019058.8
: DE 10019173.8
: DE 10032529.7
: DE 10043826.1
: PRIOR FILING DATE: 2001-04-06
: 2000-04-06
: 2000-04-07
: 2000-06-30
: 2000-09-01
: NUMBER OF SEQ ID NOS: 228
: SEQ ID NO 52
: LENGTH: 9539
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-52

Query Match 5.0%; Score 83.8; DB 14; Length 9539;
Best Local Similarity 43.0%; Pred. No. 3.2e-05;
Matches 463; Conservative 0; Mismatches 612; Indels 2; Gaps 1;

QY 4 TCACACAGCAAGCAATGAAATAGATTAATTTTCGAGGATTAAGCGAGCTAA 63
DB 1116 TCTTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1057

[illegible]

RESULT 2
US-10-239-676-198/c
; Sequence 198, Application US/10239676
; Publication No. US20030082609A1

```

GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIPENBROCK, Christian
APPLICANT: BERLIN, Rolf
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
FILE REFERENCE: 5013.1003
CURRENT APPLICATION NUMBER: US/10/239,676
CURRENT FILING DATE: 2002-09-24
PRIORITY APPLICATION NUMBER: PCT/EP01/03968
DE 10019058.8
DE 10019173.8
DE 10032529.7
DE 10043826.1
PRIORITY FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 198
LENGTH: 7823
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-198

```

[illegible]

Db	491	CTAAATAAAAAAAAAAAAAATTTAACAATTTCATATAAAAAAAAAAACTTCTAAAAAAATT	432
OY	662	ACCGGTGTGTTTGTAGAAATGACGATGAAAGAAATGATGAAAAATATCAACTGATGT	721
Db	431	TTCTCATATATAAAATCTTAAAAAATTTTAAAAAATATATAATATCAAAATACGTTA	372
OY	722	TATGAGAAACACGACGTCATATATGAGAAATTAATTCAAAAATATAAGTTGAGCTAGTAA	781
Db	371	CTTCAAAATTAACCTCAAACTTAATATTTAAAAAACAACAAAATATAAAAAAACCTATTA	312
OY	782	ATCGATGATCTATATAACGGACATTAATAGACTGGAACCTGAAATAGTTGATTAAC	841
Db	311	ATATATAAAATTCACATAAACACATAAACAATAATATAAAATATTAACAAATATAAAA	252
OY	842	AAGAAGTCCCTTTCGTTTTTGTATTTGTGAGAGATTTCAGCACTGGAGAAAAAGCTGA	901
Db	251	TAAAAACCTTTTAAAAACAATAAAAAACAATAATCATTTTAAAAACAAAAAAACCTATAT	192
OY	902	TAAAGCAAGAGATATAAAGATTTCGATACAAATAAAGCAAAAACCTAAGAGCTTTTGA	961
Db	191	TAAAAACACAAAAAAACACTACCCCTCTCTTAAAAATTAATAAAAAAAACCTTA	132
OY	962	ATTGAGTTTGTTTTGATAGGATTTTGTAGGAAATATGAAATTCATGAGATGATAT	1021
Db	131	ATAAAATTTTATATCCCTTAATATATTT-----AAATTCATCATATACCTTAACCTAA	79
OY	1022	TTTCACACAAAAAAGACCTCGACAGTAAAGAAAAAAT	1063
Db	78	TTCTAAAAAATAAAAAATATATATACGTAAACCTATTAACCAAT	37

```

RESULT 3
US-10-172-086-17/c
: Sequence 17, Application US/10172086
: Publication No. US20030113750A1
: GENERAL INFORMATION:
: APPLICANT: Epigenomics AG
: TITLE OF INVENTION: Method and nucleic acids for the differentiation
: TITLE OF INVENTION: of prostate tumors
: FILE REFERENCE:
: CURRENT APPLICATION NUMBER: US/10/172,086
: CURRENT FILING DATE: 2002-06-13
: NUMBER OF SEQ ID NOS: 116
: SEQ ID NO 17
: LENGTH: 18997
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
: US-10-172-086-17

```

[illegible]

Dp	9637	CTAAAAATAAAAAATTTTAAAAACAATCAAAACAAAACAAATAATTAAACATATAAA	9568
Qy	307	AATTGCATTAATTGAANAATACGTAGTATATCAATAATAGGGGTTTGCTATTTATTTTG	366
Dp	9567	ATATTCAAAAAATTAACAATAATACCTAATCTATTATAAAAAACAATAATAAAAAACA	9508
Qy	367	CGAAGATTGAAAAATCTG---AGTGAAGAAGAAATAGATTGGGACAGCAAAAAACCGTTG	422
Dp	9507	CAAAAAATTAACACTACAAACACGAAAAAAACCAATCATATAAAAAACCTATATTTCAAA	9448
Qy	423	CCGTTTTTTTCAAAATGACTTTGGAAAAAATTCATTTGTGAGGCGTAGGAAAATTGAAAT	482
Dp	9447	AAAAAATTTTAAATTTTACTTTAAAAAACAACAAATTTATACATAAAAATTAACCAATCAAT	9388
Qy	483	TTTTTTA--CATTGAATTTGCAAAAAATTAAGGCAAAAAAACTCAATGGAAAAAATATT	540
Dp	9387	TTACAAATTCAAATCAAACTTACAAACACAAATAAAAAATTAACCTAAAAACCGAA	9328
Qy	541	ATTATAAAAAAGAGATCGGATGTGATTTTAAACAGAAACCTGACATTGAATGAA	600
Dp	9327	AAACCAATAAAAAATTTACATATATCTCTATAAAAAATATTAAAAAACTTAACAAAA	9268
Qy	601	AAAAAGATTTGGAAAAATCTATGCTGAGAGTAAATTAAGCAAAAAAATTGGAACTC	660
Dp	9267	ATTAATAAAATTAACATATATATAAAAAATATTAAAAACAACCAATCAACACATAATA	9208
Qy	661	AAACCGGCTGTTTTGAAATGACGATGGAAGAAGATGATAAAAAATATCAACCTCGATG	720
Dp	9207	ACAAACTCTATCTTCGAAATACAAATTAATAAAAAATATATTAAAAATCCAAATCTA	9148
Qy	721	TTAATGAAGAAACGACAGTCATATATAGGAATTAATTCAAAAATTAAGTTGAGCATGTA	780
Dp	9147	ATAATATATTCCAAAAAATAATACTAAACCCCAATCCACAAACCTA-ATTAATCTCCGTA	9089
Qy	781	AATAGATGATCTAGTAACGGACTATTAGAGTGTGAACTCGAAATATGTTTGATATAA	840
Dp	9088	AAAAAATAAACTTAACCTAAACCTCTTACCTTAACAATTCAAACATTAATTTTAAATTTA	9029
Qy	841	CAGAAGTGCCTTTGTTTTGTAATTTGTGAGAAATTCAGCACTAGAAAGAAACGCTG	900
Dp	9028	TAACTATAAAAAAACCAAACTATTTTTTAAAAACAAATTAACCTATATTAATTAATAAC	8969
Qy	901	ATAATGCAAGAGAGTAAGAAGATTTGCATACATGAAGCAAAACCTAAGAGGCTTTTG	960
Dp	8968	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA	8909
Qy	961	AATTGAGCTT---GTTTTTGATAGAGATTTTGTGAGTGAAGATAGAAATCAATGAGATG	1017
Dp	8908	AAAACAATCTTAATATTTTCTACTTATTCCTCCCTTAATTAACACTTTTAATTAACACACT	8849
Qy	1018	ATATTCTCCACACAAAAAGACCTCGCAGTAAATAAACAACAAACCTAAGAACCGGACA	1077
Dp	8848	TTCAATTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA	8789
Qy	1078	CAATGAGATGATTTTAAAGGCTAAAAAATAATAATCTACATATGATC 1128	
Dp	8788	AAAACAATATCTCTCTATAAATAATCTAACTCAACAAATTAACAACTAAC 8738	

```

: RESULT 4
: US-10-239-676-51/c
: Sequence 51, Application US/10239676
: Publication No. US20030082609A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIPENBROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
: FILE REFERENCE: 5013.1003
: CURRENT APPLICATION NUMBER: US/10/239.676
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: PCT/EP01/03968
: DE 10019058.8

```

DE 10019173.8
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 51
LENGTH: 9539
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-51

Query Match 4.3%; Score 71; DB 14; Length 9539;
Best Local Similarity 43.6%; Pred. No. 0.0052;
Matches 469; Conservative 0; Mismatches 595; Indels 11; Gaps 3;

QY 7 ACACAGGCAATTTGAAAAATAGTAAATTTTCGCGAGTATTAAAGCCGACTTAAC 66
DB 6229 ATAAAAATTAATTTAAAAATCAATATTAATTAATCAAAAAATTTAAATTAAC 6170
QY 67 AAATGAGTGAAGAGAAAAAGAAAAATTAATTAATTTGAGTTAGTAAAGAGAAAG 126
DB 6169 ATAAATTTTAAAAAACAATATTAACGAATATTAACAAAAACGAAGCTTAAAT 6110
QY 127 AAAAAATTAAGAACCTCGGCTTAACAGTCGAAAAACCAAAATTAATAAAGAAAGA 186
DB 6109 TCAAAAACACATTTTAACTCAAAAATTAATTAATAAACAACGAATTTAAAAATCGA 6050
QY 187 GACTGTGATTTTATGGAATCGAGGAAAGAAAT-TTAAATTTTCATTTTCGAGG 245
DB 6049 AACGAATTTTAAATCAAAAATTAATTAATAAATTAATTAATTAATTAATTAAT 5990
QY 246 GATTATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 305
DB 5989 TCATTAATAAATCTAAAAATTTCTAAAAATTTCTAAAAATTTCTAAAAATTC 5930
QY 306 GAATTTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365
DB 5929 AAAAAATCTATTTAAACCAAAATTAATTAATTAATAAATAAATAAATAAATAA 5870
QY 366 GCGAAGATTTGAAATCTGAGTGAAGAAATAGTTGCGAGAGCAAAAAACCTTGGCG 425
DB 5869 AAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5810
QY 426 TTTTTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 485
DB 5809 TAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5750
QY 486 TTACATTTGATTTGAAATTAAGCAAAAGAAACCAATGAAATTAATTAATTAAT 545
DB 5749 TAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5691
QY 546 AAAAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605
DB 5690 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5631
QY 606 GATTGGAATAATCTATGCTGAGAGTGAATTAAGCAAAAAATTTGGAGCTCAACC 665
DB 5630 AAAAAACAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5571
QY 666 GGTGTGTTTGAATTAAGCATGAAGAAATGATGAAATATCAACCTGATGATTAAT 725
DB 5570 AAAAACTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5511
QY 726 GAAGAAGCAGAGTCATATAGAAATTTTCAAAATTAAGTTGACATAGTAAATCA 785
DB 5510 ACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5451

QY 786 GATGATCTAGTAACGGGACTATTAAGAGTGGAACTCGAAATAGTTTGAATTAACAGA 845
DB 5450 AATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5391
QY 846 AGTGCCTTTCCTTTTGTATTTTGAGAGAAATTCAGCACTGAGAAAAAGAGCTGATAT 905
DB 5390 -----TTAATTTTAAAAAACAACCTTAAAAATTAACGAAATTAATTAATTAATTAAT 5340
QY 906 GCAAGAGAGTAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 965
DB 5339 TCCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5280
QY 966 AGTTTGTGTTTGAATTAAGGATTTTGTGAGTGAATTAATTAATTAATTAATTAATTAAT 1025
DB 5279 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5220
QY 1026 CACAACAAAAAGACTCTCAAGTAAAGAAAAACAATGAAGAAGAGGACACAA 1080
DB 5219 AAAAAATCAAACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTAA 5165

RESULT 5
US-09-960-352-14521

Sequence 14521, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengping

APPLICANT: Byatt, John C.

APPLICANT: Mathiasagan, Nagappan

TITLE OF INVENTION: NOCTIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960.352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 14521

LENGTH: 411

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 62-LIB3058-002-Q1-K1-H6

US-09-960-352-14521

Query Match 4.2%; Score 70; DB 11; Length 411;
Best Local Similarity 54.7%; Pred. No. 0.0037;
Matches 139; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 62 AAAAAATGAGTGAAGAGAAAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 121
DB 117 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 176
QY 122 GAAGAAATTAATTAAGAGAGCTCGCTTAACAGTCGAAACAGAAATTAATTAATTAAT 181
DB 177 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 236
QY 242 GAGGATTAATTTGTTGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 301
DB 297 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 356
QY 302 TGTGATTTGACA 315
DB 357 TAAAAATTTTAA 370

RESULT 6

US-09-960-352-3400/C

Sequence 3400, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengding
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 3400
LENGTH: 446
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 15-LIB3058-052-Q1-K1-D11
US-09-960-352-3400

Query Match
Best Local Similarity 4.2%; Score 69.4; DB 11; Length 446;
Matches 193; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 369 AAGATTGAAATCTGAGTGAAGAAATAGTTGGCAGAGCAAAAACCCCTTGCCGTTT 428
DB 421 AAAAATTAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 362
QY 429 TTTTCAATGACTTTGGAATAATTCATGTGAGCGGTAGCAACCTTGAAATTTT 488
DB 361 TATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 302
QY 489 CATTGGAATTTGAAATAATAAGCAAACTCAATGGAATAATTTATTATATA 548
DB 301 AAAAATTAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 242
QY 549 AAAAGAGATCGATATGATTTTAAAGCAGAAATCTGACCTTGATGAAAAAAGAT 608
DB 241 AAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 182
QY 609 TTGGAATAATCTATCTGAGAGTGAATTAAGCAAAATTTGGAACTCAACCCGT 668
DB 181 AAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 122
QY 669 GTTGTAGAAATGACGATGAAGAAATGATGAAATATCACTCGATTTATATA 728
DB 121 AAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 62
QY 729 GAAACGACGTCATATATGAAATTTATTCATAAATAA 767
DB 61 AAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 23

RESULT 7
US-09-878-574-4653/c
Sequence 4653, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 4653
LENGTH: 821
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)...(821)

OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: LIB3028-039-Q1-B1-F4
US-09-878-574-4653

Query Match
Best Local Similarity 4.2%; Score 69; DB 11; Length 821;
Matches 270; Conservative 0; Mismatches 340; Indels 0; Gaps 0;

QY 6 TACACAGGCAATTTGAAAAATAGTAAATTTTCGCGATTTAAAGCCCACTTAAA 65
DB 821 TATAGAGCAGAGAAAGATATATAGTTTATGATTTTGTAAAGAAAGAAATAGTA 762
QY 66 CAATGAGTGAAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 125
DB 761 GTTAAATATATTAAGAGAAATTTAAATTAATTAATTAATTAATTAATTAATTA 702
QY 126 GAAAAATTAAGAAAGACCTCGCTTACAGTGCAGAAAAACGAAATTAATTAAGAAAG 185
DB 701 AAAAGATTAAGAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 642
QY 186 ACACCTGCTTTTATGAAATCTGAGGAAAGAAATTTATTTTTCATTTTCGAGC 245
DB 641 GAAAGTAAATGATATACAAATAGAAACAAAGAAAGATTAATTAATTAATTAATTA 582
QY 246 GATTAAATTTGTTAAGTATGATGAAATAATCTAGATTAATAATTAATTAATTAATTA 305
DB 581 AAAGAAAGTGTCTTGAATACAAATAATTAATTAATTAATTAATTAATTAATTAATTA 522
QY 306 GAATTTGACATTTTGAATACGATATATCAATTAATTAATTAATTAATTAATTAATTA 365
DB 521 AAATTAATTAAGAGAAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 462
QY 366 GCGAAGATGAAATCTGAGTGAAGAAATAGTTTGCAGAGCAAAAAACCCCTTCCG 425
DB 461 TATGTAAATATTTTCTGCAAAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 402
QY 426 TTTTTCATGACTTTGGAATAATTTATTTGTCGCGGATTCGAAATTTT 485
DB 401 TTAGTTTATGTGATATGATTTATATGATTTATGATTTATGATTTATGATTTATGATTT 342
QY 486 TTACATTTGAAATTTGAAAAATTAAGCAAAAGAAATCAATTAATTAATTAATTAATTA 545
DB 341 GATTCTTAATAATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 282
QY 546 AAAAAGAGATCGATTTGATTTTAAACAGAAATCTGACATTTGATTAATTAATTA 605
DB 281 AAAAATGAGATGAAAGTTTGAATTTTAAAGAAATTAATTAATTAATTAATTAATTA 222
QY 606 GATTGGA 615
DB 221 AATTGAGAA 212

RESULT 8
US-10-198-846-6381/c
Sequence 6381, Application US/10198846
Publication No. US20030099974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
FILE REFERENCE: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6381

```

: LENGTH: 960
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1, 8, 66, 71, 79, 81, 84, 85, 86, 88, 117, 119, 125, 134,
: LOCATION: 149, 156, 162, 166, 172, 179, 180, 194, 230, 249, 250, 253,
: LOCATION: 268, 271, 273, 274, 275, 276, 277, 286, 291, 299, 312, 333,
: LOCATION: 359, 366, 374, 393, 404, 406, 409, 411, 422, 424, 427
: OTHER INFORMATION: n = A,T,C or G
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 428, 433, 450, 465, 466, 468, 471, 487, 510, 512, 513, 516,
: LOCATION: 517, 518, 528, 530, 534, 539, 553, 555, 559, 575, 580, 588,
: LOCATION: 590, 610, 625, 626, 642, 645, 656, 658, 659, 675, 685, 686,
: LOCATION: 690, 700, 702, 706, 716, 731, 745, 748, 756, 798, 802
: OTHER INFORMATION: n = A,T,C or G
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 807, 808, 819, 820, 821, 830, 838, 840, 844, 845, 849, 861,
: LOCATION: 871, 874, 875, 891, 892, 895, 898, 900, 901, 902, 904, 907,
: LOCATION: 917, 923, 925, 926, 927, 929, 955
: OTHER INFORMATION: n = A,T,C or G
: US-10-198-846-6381

```

```

Query Match      4.2%; Score 69; DB 14; Length 960;
Best Local Similarity 39.8%; Pred. No. 0.0067;
Matches 317; Conservative 0; Mismatches 468; Indels 12; Gaps 2;

```

```

OY 22 AAAAAATGATTAATTTTCGAGGTATTAAGCCGACTTAACCAATGAGTGAAG 81
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 951 AAAAAAATTTTGTGTTNNNNNACCCTTAACTTTTAAATNTNNNNNTTAAAN 892
OY 82 AAAAAAATTAATCATATTTTGTGTTAAAGAGAAGAAAAATTAAGAG 141
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 891 NAAAAAATTTTAAANNNTAATTTTNNAAAACTAATTTTAAATNTAATNTAATTTT 832
OY 142 ACCCTGGCTTAACCTCGAAGAACAGAAATATTAAGAGAGACTGTGATTTTAA 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 831 TTTTAAAAAANNAAAAATTAANNNTTTTNTAANTTATTAACCAAAATTTTAAAA 772
OY 202 TCGAATTCGTGAGGAAAAAATTTTATTTTCATTTTCGAGGATTAATTTGTTAA 261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 771 AAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 712
OY 262 GTTGTGAAAAATCTAGATTAATAAATGACATCAAAATGTGTAATTTGACATTAATG 321
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 711 TTTAANCANNTTTTNTTTTNTTNNAAAAAANNTTTTAAATTAATAAANNTTTT 652
OY 322 AAATCGTAGTATCAATATGAGGCTTGTCTATTTTATTTTTCGAGGATTTGAATTC 381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 651 TTTTNTNANATATTTTAAANAAAAAANNTTTTNTATNTATTAATAAATTTTAAAT 592
OY 382 TGAGTGAAGAAATAGTTTGGGAGAGCAAAACCCCTGCGTTTTCGAATGACT 441
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 591 TTTNT-----TTTCTNAAAAAANAAAAATTTTNTAANNTTTTAAACCT 541
OY 442 TTGAAAAAATTCATTTGAGCGGTAGAGCAAACTTTTGAATTTTACATTGGAATTTG 501
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 540 TTAATTTAANANATTTTNTTNNNTTNNANATTAATAATTTATTTTNTTANNTCTAT 481
OY 502 AAAAAATTAAGCAAAAGAACTCAATGCAAAAAATTTATTTATTAATAAAGAGATCGG 561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 480 AAAAAAANANANANANATTAATAATTTTNNAAAAAATTTTATNTTANNTAANNAANA 421
OY 562 ATATGATTTTAAAGCAAGAAACGACATTTGAATGAAGAAAAAATTTGAA---AAA 618
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 420 AATTTAANNTAANAAAAAATTTTAAATTAATAAATTAATTAATTTAAACNTCAAT 361
OY 619 TCTATGCTGAGAGTGAATTAAGCAAAAAAATTTGGAACCTCAACCCGGTGTGTTTAA 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 360 TTTATTTTATTAATTAATAAATAATTTTNTTATTAATAAATAATTTTATTTAATAA 301

```

```

OY 679 AATGACGATGAAGAAATGATGAAATATCAACCTCGATTTAATGACAACAGCAG 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 300 AAAAAAATTAATTTATTAAGNNNNNTNTTAAAAAATTAATNTAATNTAATTTA 241
OY 739 GTCAATTAAGAAATTTATCAAAAAATTAAGTGAATGATTAACAGATGATCTAGTAA 798
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 240 ATTTAATTTGNNAAAAAATTTAATAAATTTATTTTATTAACNATGTTATTAATAA 181
OY 799 CGGACTATTAGAGTGT 815
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 180 NNTTAAANNAATGAGNAT 164

```

```

RESULT 9
US-09-960-352-4630
: Sequence 4630; Application US/09960352
: Patent No. US20020137139A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengbing
: APPLICANT: Byatt, John C.
: APPLICANT: Mathalagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: FILE REFERENCE: 16511.006/37-21(10298)C
: CURRENT APPLICATION NUMBER: US/09/960,352
: CURRENT FILING DATE: 2001-09-24
: NUMBER OF SEQ ID NOS: 1512
: SEQ ID NO 4630
: LENGTH: 327
: TYPE: DNA
: ORGANISM: Bos taurus
: OTHER INFORMATION: Clone ID: 20-LIB3058-048-Q1-K1-E11
: US-09-960-352-4630

```

```

Query Match      4.1%; Score 67.4; DB 11; Length 327;
Best Local Similarity 51.9%; Pred. No. 0.0097;
Matches 152; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

OY 17 AATGAAAAATGATTAATTTTCGAGGTATTAAGCCGACTTAACCAATGAGTGA 76
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 32 AAGTAAAGTTTTCATGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 91
OY 77 AGAAGAAAAAAGAAAAATTAATTAATTAATTTTGAATTAAGAGAAAAAATTA 136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 92 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 151
OY 137 AGAAGCTCGGCTTAACAGTGAAGAACCAAGAAATTAATAAAGAGAGCTGTGAT 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 152 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 211
OY 197 TTTAATGAAATCTGAGGAAAAAATTTTAAATTTTCAATTTTCGAGGATTAATTTGT 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 212 AGGAGAGAAAGTTTAAAGAAAAAATTTTAAAGAAATTTTAAAGTAAATTTT 271
OY 257 TGTAAAGTGAAGAAAAATCTAGATTAATAAATGACATCAAAAAATGTGTTGAAT 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 272 TAAAGGAAAAAATTTTAAAAAATTTTAAAAAAGGGGGAATTTTAAAAAATTTGAAT 324

```

```

RESULT 10
US-10-239-676-185/c
: Sequence 185; Application US/10239676
: Publication No. US20030082609A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIEPENBROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
: FILE REFERENCE: 5013.1003
: CURRENT APPLICATION NUMBER: US/10/239,676
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: PCT/EP01/03968

```

```
DE 10019058.8
DE 10019173.8
DE 10033529.7
DE 10043826.1
PRIOR FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 185
LENGTH: 7657
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-185
```

```
Query Match
Best Local Similarity 47.0%; Score 67.4; DB 14; Length 7657;
Matches 209; Conservative 0; Mismatches 236; Indels 0; Gaps 0;
```

```
QY 494 GAAATTTGAAAAATAGGCAAAAGAACTCAATGAAAAATATTATTATTAATAAAG 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5979 GAACATTAAATATATCACCACAAAAAGTAACGAAATATAAAAATTAATAAATA 5920
QY 554 GAGATCGATATGATTTTAAAGCAGAAACTGACATTCATGAAAAAAGATTGGA 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5919 AAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 5860
QY 614 AAAAACTATGCTGAGATGATTAATTAAGCAAAAAAATTTGGACATCAACCCGGTGTGT 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5859 TAATTAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5800
QY 674 TTGAATGAGCAGTAAAGAAATGATGAAAAATATCAACCTGATGTTAATGAGAAC 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5799 AAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 5740
QY 734 AGAGAGTCAATATAGAAATTAATTCAAAAATTAAGTGAACATAGTAATCATGATCT 793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5739 AAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 5680
QY 794 AGTAACGAGCTTTAGAGTGTGAGTGAACCTGAAATAGTTTGTGTAACAAGAGTCCCT 853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5679 AAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 5620
QY 854 TCGTTTTGTATTGTGAGAGAAATTCAGCAACTGAGAAAAAGAGCTGATATGCAAGAG 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5619 TAAATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5560
QY 914 AGTAAAGATTCGATCAATGAAA 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5559 ATAAAAATTAATAATAAACCAATA 5535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11
US-09-960-352-573/C

Sequence 573, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathalagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 573
LENGTH: 428
TYPE: DNA

ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 03-LIB3058-031-Q1-K1-A11
US-09-960-352-573

Query Match
Best Local Similarity 52.3%; Score 66.6; DB 11; Length 428;
Matches 147; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```
QY 502 AAAAAATAGGCAAAAGAACTCAATGAAAAATATTATTATTAATAAAGAGATCGG 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 390 AAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 331
QY 562 ATATGATTTTAAAGCAGAAACCTGACATGTAAGTAAAAAAGATTGGAATAATCT 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 330 AAAAAAGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 271
QY 622 ATGCTGAGATGTAATTAAGCAAAAAATTTGGAACTGACCCGGTGTGTTAGAAA 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 270 AAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 211
QY 682 TGACGATGAAAGAAATGATGAAAAATATCACTCATGTTATGAAAGACGACGTC 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 210 AAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 151
QY 742 AATATAGAAATTAATTCAAAAATTAAGTTGAGCATAGTAA 782
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 150 AATATATATATATATATATATATATATATATATATATATATATATATATATAC 110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 12

US-10-239-676-35/C
Sequence 35, Application US/10239676
Publication No. US20030082609A1

GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
CURRENT APPLICATION NUMBER: US/10/239,676
CURRENT FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: PCT/EP01/03968

DE 10019058.8
DE 10019173.8
DE 10033529.7
DE 10043826.1
PRIOR FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 35
LENGTH: 12405
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:
NAME/KEY: unsure
LOCATION: (7895)
US-10-239-676-35

Query Match
Best Local Similarity 42.8%; Score 65.4; DB 14; Length 12405;
Matches 458; Conservative 0; Mismatches 596; Indels 15; Gaps 2;

```
QY 63 AACCAATGCTGTAAGAGAGAAAGAAAAATATATATATATATATATATATATATAG 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 9523 AAAAAAACAACAACAAGCAAAATTCATAAATATATATATATATATATATATATAT 9464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 123 AAGAGAAATTAATAAGACAGCTCGCTT-----ACAGTCGAAAAACCAAGAAATATATA 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db	9463	AAACAAACAAACCAATATACCAAAATATATAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAC	9406
OY	179	AAGAAAGAGACTGATGATTTTAAATCGAANAATCGTAGGAAAGAAATTTTAAATTTTCAT	238
Db	9403	AAAAAAAAAAAAACCAACAAATACCATATCAATTAATCCATACCAAAAAAAAAATAAATACCT	9344
OY	239	TTGAGGAGATTAATTTGTTGTAAGTGTAGTGAAGAAATCTAGATTAATAAATGCGATCAAA	298
Db	9343	AAATTAATAAATTAACCTATATAAATAATATAAATAATTTTAACATCTTAATATATTTTAAA	9287
OY	299	ATGCTGAATTTGCATTTATTTGAAATACGATATATCAATATATGGGGTTTGTCTATT	358
Db	9283	ATATACAAACAAACAAATCCCTAAAAAAAATCAAAAATTAACCTCAATTTTAAATAAAA	9222
OY	359	TTATTTTCGGAAGATTGAAATCTGAGTGAAAGAAATATGTTGGCAGACCAAAAAAC	418
Db	9223	CAATTCACAAATTAATTAATATCTACAAAAAAAACCAATTTTAAACATTAATAAATAAAAAC	9166
OY	419	CTTGCCGCTTTTTCACAAATGACTTTGGAAAAAATTCATTTGAGCGGTAGCGAACTTGG	478
Db	9163	TAACCTTAACCTTCTAACATCCACAAAAAAAACCAACACCTAACACGAAATATTA	9104
OY	479	AAATTTTTCACATGGAATTTGAAAAAATTAAGCAAAAGAAACCTCAATGGAAAAATA	538
Db	9103	TACTTAAAAAAAATCCAACTAAAAATATTAATTCATATATCTTCAACGATTAACAA	9044
OY	539	TTATTTATAAAAAAGAGATCGGATATGATTTTAAAGCGAAAACTGACATTTGAATGA	598
Db	9043	TTTATCCCAAAATTAATTAATCCACTATTTTAACATTAATAAATAATTAACAT-----A	8995
-OY	599	AAAAAAGATTTGAAAAAATCTATGCTGAGAGTGAATTAAGCAAAAAAATTTGGAAAC	658
Db	8994	TAAAAAATAATTAACCAACACTCTAAAAAATTTAAATCCAAAAAATTAATAAACAAC	8933
OY	659	TCAACCGGCTTGTGTTTGAATGAGATGAAAGAAATGATGAAAAATATCAACCTCGA	718
Db	8934	CTACCACCTAAATCAAAAAAATCCAAAAAATATTTAAATCTTAAATCCAAAAAATAAA	8875
OY	719	TGTTAATGAAGAAACAGCGCTCAATATAGGAAATTTCAAAAAATAAGTTGACATAG	778
Db	8874	ATATATAAAAAATATTAACCATATCAATTAATTACTATCAAAACAAAAATTTAAAACTAATC	8815
OY	779	TAAATCAGATGATCTAGTAACGGGCACTTATAGAGTGTGAAGCTGAAATAGTTTGTATTA	838
Db	8814	AATATATTAACCTTTAAAAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA	8755
OY	839	AACAAAGAGTCCCTTTGTTTGTATTTGTAGAGAAATTCAGCAACTGAGAAAAAGAC	898
Db	8754	TAAACCTAAATCTTAATTAATAAATAATTTCAAAAAAATAAATAAATAAATAAATAAATAA	8635
OY	899	TGATTAATGCAAGAGATTAAGAAATTCGATTCGAATGAAGCAAAAACTTAAGAGCTTT	958
Db	8694	CAACCTTAATAAATAATCTTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA	8635
OY	959	TGAATTTGAGTTTGTGTTTGAATGAATTTTGTGTAATAATGAATTCATGGAATGA	1018
Db	8634	ACTATTAATCTCCAAAAAATTTTGTTCCTTTTAAATTAATAAATTAATAAATTAATTAAT	8575
OY	1019	TATTTTCACAAAAAAGACTCTGCAATTAAGAAAAAACAATGAAGAAGCGGACAC	1078
Db	8574	ATTAATATATCTAAATAAATAATATCTAATTAATAAATAAATAAATAAATAAATAAATAA	8515
OY	1079	AATGAGATATTTTAAAGAGCTAAATAAATAATTAATAATCTACATATGAT	1127
Db	8514	AAAAAAAAAATAGCTATAAATAATACATTAATTAATAAATAAATAAATAAATAAATAAATAA	8466

RESULT 13
US-10-094-240-10/C
; Sequence 10, Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.

```

: TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
: FILE REFERENCE: N8289
: CURRENT APPLICATION NUMBER: US/10/094,240
: CURRENT FILING DATE: 2001-03-08
: PRIOR APPLICATION NUMBER: 10/056,405
: PRIOR FILING DATE: 2002-01-24
: PRIOR APPLICATION NUMBER: 60/264,649
: PRIOR FILING DATE: 2001-01-26
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 10
: LENGTH: 4985
: TYPE: DNA
: ORGANISM: Anopheles gambiae
US-10-094-240-10

```

Query Match	3.9%	Score 64.8	DB 14	Length 4985
Best Local Similarity	45.1%	Pred. No. 0.053		
Matches	278	Conservative	0	Mismatches 337
			IndeIs	1
			Gaps	1
OY	504	AAATAGCGCAAAAGAACTCAAAATGAGAAAAATATTTTATATAAAAAAGAGATCGAT	563	
Db	2967	AATATATAGAGATTAATTAATACAAATGAGCATATATATATATATATTAATTAATTAAT	2908	
OY	564	ATGGATTTTAAAGCGAGAAACGTGACATGATGAAAAAAGATTGGAAAAATCTAT	623	
Db	2907	AATATATATTAATTAATAAAAAAATTAATTTATATAGAAATGTAATTAATTAATTAATTA	2848	
OY	624	GCTGAGACTGAATTAAGACAAAAAATTTGGAACTCAACCCGGTGTGTTTAGAAATG	683	
Db	2847	TATTAATAATATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2789	
OY	684	ACGATGGAAGAAAGATGAAAAAATATCAACCTCGATGTTAATGGAAGAACGCGGTCAA	743	
Db	2788	ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2728	
OY	744	TATAGGAATTAATTCAAAAATTAAGTTGAGCATAGTAATCATGATGCTAGTAACGGGA	803	
Db	2728	TATGTAATAATTAATGAATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2668	
OY	804	CTATTAGAGTGGACCTCGAAATAGTTTGATTAACAAAGAGCCTTTCGTTTTGT	863	
Db	2668	ATTAATTAATTAATTAATTAATTAATTTATTTGTTATTAATTAATTAATTAATTAATTAAT	2609	
OY	864	ATTTGCGAGAAATTCAGCAACTGAGAAAGACGTGATTCGAAGAAGAGTAAAGAT	923	
Db	2608	AATATATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2549	
OY	924	TTTCGATACATGAAGAAGCAAAACCTTAAGAGCCTTTGATTGAGTGTGTTTGGATGAG	983	
Db	2548	AATATATATTAATTAATTAAGTAATTAATTTATTAATTAATTAATTAATTAATTAATTAAT	2489	
OY	984	GATTTTGGAGTGAATTAAGAAATTCAAATGGAATGATTTTTCACAACAAAAAGACTCT	1043	
Db	2488	AATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2429	
OY	1044	GCAAGTAAAGAAAAACAATGAAGAAGCGGACCAATGATGATTAATTTTAAAGGCTA	1103	
Db	2428	AATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2369	
OY	1104	AAAAATATAATTAATCTA 1119		
Db	2368	AATATATATTAATTAATTA 2353		

RESULT 14
US-10-239-676-90/c
; Sequence 90, Application US/10239676/
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt


```
OY 538 ATTATTTAAAAAGGAGATCGATATGATTTTAAAAAGCAAACTGACATTGAATG 597
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4658 AATTATATTCCTATATAAAAAATAATACAACTAATTAATCTAAAAAACCCAAAAAAA 4539
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 598 AAAAAA---GATTTGAAAAATCTATGCTGAGAGTGAATTAAGCAAAAAAATTG 653
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4598 AAAAAATTAATCTACCTTAATAAAATCTTAATAATTAACCTTAATAAAATTAATTAA 4539
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 654 GGAACTCACCCGGTGTGTTTGTAGAAATGAGATGAAGAATGATGAAAAATATCAAC 713
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4538 TAAATTTTACCTAATTAATTAATAAAAAAACAACAAAAACACCAAAAAAT 4479
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 714 CTGATGTAAATGAAGAACAGCGTCATATAGGAATTTCAAAAATTAAGTTGAG 773
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4478 ATATAAAAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4419
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 774 CATAGTAATCAGATGATCTAGTAACGGGACTATTAGAGTGTGAACTCGAAATAGTTT 833
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4418 AATTAATAATTCATCTATTAAATTAACCAACATTAATTAATTAATTAATTAATTA 4359
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 834 GATTAACACAGAGT 848
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4358 TCAAAAAAAATTT 4344
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: July 17, 2003, 07:00:51
Job time : 632.027 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2003, 02:19:36 ; Search time 2699.14 Seconds
(without alignments)
11082.453 Million cell updates/sec

Title: US-08-876-132-2
Perfect score: 1847
Sequence: 1 ACATCTCAACACAGTTTAA...AACACAGCAAAAGAGATCT 1847

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107.8	5.8	1101	17	CNS00EVL
2	101.2	5.5	1101	17	CNS00EVL
3	100.6	5.4	1135	17	CNS0336Q
4	97.6	5.3	1201	17	CNS0167M
5	97	5.3	974	17	CNS0017T
6	96.2	5.2	1029	17	CNS01ZGM

7	94.4	5.1	829	17	CNS03LUB
8	93.6	5.1	1025	17	CNS01472
9	93.4	5.1	1101	17	CNS003BD
10	93.2	5.0	1101	17	CNS0039G
11	93	5.0	660	17	BH183498
12	92.2	5.0	660	17	CNS070NJ
13	92.2	5.0	1036	17	CNS03LWJ
14	91.8	5.0	804	17	B12681
15	91.8	5.0	1885	17	BE420745
16	91.6	5.0	934	13	BM415088
17	91.6	5.0	1101	17	CNS0039R
18	91	4.9	1101	17	CNS014XJ
19	91	4.9	1187	17	B11102
20	90.4	4.9	942	17	CNS018G5
21	90.2	4.9	982	13	BM415348
22	90.2	4.9	960	17	AG031602
23	89.8	4.9	782	10	BB622182
24	89.6	4.9	1101	17	CNS0021J
25	88.2	4.8	1101	17	CNS0005X
26	88	4.8	1101	17	CNS00E07
27	88	4.8	1225	17	CNS0161D
28	87.8	4.8	879	17	CNS01JRG
29	87.8	4.8	1139	17	AO897537
30	87	4.7	714	17	AO271175
31	87	4.7	910	13	BM415636
32	86.8	4.7	1101	17	CNS012JN
33	86.6	4.7	668	9	AL514901
34	86.6	4.7	907	17	CNS02134
35	86	4.7	795	17	AG044101
36	85.6	4.6	1059	17	CNS0022B
37	85.4	4.6	994	17	CNS04N0J
38	85.2	4.6	990	17	CNS0060T
39	85	4.6	1036	17	CNS00599
40	84.2	4.6	890	17	AQ026918
41	84.2	4.6	915	12	BG520365
42	83.8	4.5	714	17	AQ050116
43	83.6	4.5	948	13	BM415494
44	83.6	4.5	1190	17	CNS020N7
45	83.4	4.5	876	17	CNS009G1

ALIGNMENTS

RESULT 1
CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL069706.1 GI:4949849
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial

[illegible]

OY		948	AATTTTCGAGCAAAAATATTTTTGAAFTTTTTTAAGCGCATCTTCCACGACTTT	1007
Dd		242	AAA	301
OY		1008	TGCCATATTTTAAACCCTGACTATCTTTTAAAGTTATATGATATATCCTTGATATATAA	1067
Dd		302	AAA	361
OY		1068	GTAATGTAAAAACGAGTAAAAACAATACTTATATATTAAATTCGAATTAATTGACA	1127
Dd		362	AAA	421
OY		1128	GTGATTAATTTATATATTAAGAGATATATCTATTACCTTAATATATACTAAAAAAGG	1187
Dd		422	AAA	481
OY		1188	TAAATATATGTGATGTGTATTTAAAAAAGCAFTTAGAAATGAGAATFAGAACATTTAAAAA	1247
Dd		482	AAAAAAAAAAAAAAAAAAAAAATTAAAAAATAAAAAAAAAAAAAAAAAAAAAAATT	541
OY		1248	AGACGGTGATATCAAACTCTTCTTAACAATCTTGCACTACTCTTGATATAGATAAAGCAAT	1307
Dd		542	AAAAAAAAAAAAAAAAAAAAAATTAAAAAATAAAAAAAAAAAAAAAAAAAAAAATT	601
OY		1308	AAATGTGATGATGAATGTGGCGATTAATTAATACTCAAATTAATTCATTGATGAAGTTTGA	1367
Dd		602	AAAAAAAAAAAAAAAAAATTAAAAAATAAAAAAAAAAAAAAAAAAAAAAATT	661
OY		1368	TCTTCTTGATGTGAGCACAAATTTCCGTGGGCTTCAATATAAATATACAGACGAA	1427
Dd		662	AAAAAAAAAATTTTAAAAAATAAAAAAAAAAAAAAAATTAATTAATAA	721
OY		1428	TGAATATTTATCATCAGCTAAACGTA - AATGATTTTAAAAAATCTCGTCTTATTC	1488
Dd		722	AAAAAAAAAAAAAAAAAAAAAATTACATNTAAAAAATAAAAAAAAAAAAAAAATCYAAACAAT	781
OY		1487	ATTCGATCAATGGGAGACAATTTAATTACATGAGATGATGCATATCATCACAAGATGA	1548
Dd		782	AAAAAAAAAAAAATTAATAAAYVATTAATTAATAAAAAAACACACYAATAAAAAA	841
OY		1547	TTAAAGATTTGATATATTTACACATATGAGAAATTAACATATGAAATA - TAAATATAG	1602
Dd		842	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAATAATTAATTAATTAATAA	901
OY		1604	AGAAATATTTGATTTCCGAATTAACAAAAGATATAAAGAAATCAGCTTTAAAAATATC	1663
Dd		902	AAAAACAATAAAAAACAATAAATTAATTAACCTTAACCAAAAAAAAAAAAAAAAAAAAAA	961
OY		1664	TGCATTTGATGATGTTTAAAAAGTATGATAAATAATTTATGATGAGAGATTTGCAACT	1723
Dd		962	AAAAAAAAAAAAATTAATAAATAAAAAAATAAATAAACAATTAATAAATAAATAATTAACW	1021
OY		1724	AAAGATATTTGAAATATTCGTATGAATAATCCATTTGATGATCCAGATGATGSCATTAAGAA	1783
Dd		1022	ATAAATAAAAAAAAAAAAAAAAAAATTCATAAATTAATAAATAAATAAATAAATAA	1081
OY		1784	ATCAATCAATCCACAGAACAAATGCTTATGCTTCCGATTTAAAAAACAAGCAA	1837
Dd		1082	AAAAAAAAAAAAAAAAAAAAAAAAAATTAATAAATAATTAATTAATAAATAAATACTAA	1135

RESULT 4
 CNS0167M/c
 LOCUS
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
 BRON1M24 of DrosBac library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL106396
 VERSION AL106396.1 GI:5621701
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;


```

Db      681  AAAAAATNSAAAAAABAAAAAGTGNCAKAAAAAARAAAAAATTTTTTTTTT
Qy      1219  TTGAGAAAAATGAAATAGACATTTATAAAAAGCGGTATTCGAATTTCTTACATAC
Db      621  GKTAKKRAKADAMWTKTAMATGKTKKASTSARADARKAAAAAATTTTTTTTTT
Qy      1279  TTGCATTACTTGTGATATAGATTAAGCATTAATGATGATGATGCGATATTATTAAC
Db      561  TGKKRTAKAKAGAKAKAKKDDAAAAAATATGKNT-ACRAATKABGCAAAAAAGVAA
Qy      1339  TCAAAATTTATCACTTGATGAAAGTTTGATCTTCTTGATGAGCAATTTGCGCTGG
Db      502  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1399  GCTTTCAATTAATTAATACGAGACGAAATGATTAATTAATTAATTAATTAATTAAT
Db      442  TAAAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1459  GATTTTAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      382  ADAABBRKABMAKBBBCKSKBTSTTTTDTAAAAAANBAKAAABMACWATTGT
Qy      1519  GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      322  RTANAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1579  ATTAACCATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      262  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1639  AAAAAAGTCAGTCTTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      202  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1699  TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      142  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1759  ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      82  NAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1819  GCATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
Db      22  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

```

```

TITLE      Weissbach,J.
REFERENCE  3 (bases 1 to 1029)
AUTHORS    Direct Submission
JOURNAL    Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
             scale clone-end sequencing project of the Tetraodon nigroviridis
             genome. For more information, please take a look at
             http://www.genoscope.cns.fr/tetraodon.
FEATURES   source
            1..1029
            /organism="Tetraodon nigroviridis"
            /db_xref="taxon:99883"
            /clone="220112"
            /note="Genoscope sequence ID : COAG220BE06SP1-end"
            /note="PUC-ori"
BASE COUNT 297 a 124 c 66 g 457 t 85 others
ORIGIN
Query Match 5.2%; Score 96.2; DB 17; Length 1029;
Best Local Similarity 45.6%; Pred. No. 7.5e-07;
Matches 267; Conservative 34; Mismatches 280; Indels 5; Gaps 2;

Qy      1056  TTGATATTAAGTATGATTTAAAAACGCTAAACAACTTAATTAATTAATTCGAA
Db      1022  TAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1116  TTATATTTGACGATGATTTAATATATTAAGATATATGATGATGATGATGATGATGAT
Db      962  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1176  TAAAAAAGAGTAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      902  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1235  AACATTTAAAAAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      842  AATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT
Qy      1295  TAGATTAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      782  TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT
Qy      1355  ATGAAAGTTTGAATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      726  ATTATTAATAAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
Qy      1415  TACAGACGCAATGATATTTATTCATGACCTAAACTGAAATGATTTAAAAATAC
Db      666  AAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT
Qy      1475  CGTTCTTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Db      606  TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT
Qy      1535  ATCAAGATGATTAAGATTTAGATATTTACATATGAGAAATTAACCATTTGATTA
Db      546  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
Qy      1595  TAAATTAAGCATTTATTTGATTTGCAATTAACAAAGATTTAA
Db      486  AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

```

RESULT 6
 CNS01ZGM
 LOCUS
 DEFINITION
 1029 bp DNA linear GSS 12-MAY-2000
 Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
 220112 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 AL174271 GI:7812328
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Tetraodon nigroviridis.
 Tetraodon nigroviridis.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 1029)
 Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished
 2 (bases 1 to 1029)
 Roest-Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source
 1..1029
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="220112"
 /note="Genoscope sequence ID : COAG220BE06SP1-end"
 /note="PUC-ori"
 BASE COUNT
 297 a 124 c 66 g 457 t 85 others
 ORIGIN
 Query Match 5.2%; Score 96.2; DB 17; Length 1029;
 Best Local Similarity 45.6%; Pred. No. 7.5e-07;
 Matches 267; Conservative 34; Mismatches 280; Indels 5; Gaps 2;
 Qy 1056 TTGATATTAAGTATGATTTAAAAACGCTAAACAACTTAATTAATTAATTCGAA 1115
 Db 1022 TAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 963
 Qy 1116 TTATATTTGACGATGATTTAATATATTAAGATATATGATGATGATGATGATGATGAT 1175
 Db 962 AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 903
 Qy 1176 TAAAAAAGAGTAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1234
 Db 902 AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
 Qy 1235 AACATTTAAAAAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1294
 Db 842 AATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 783
 Qy 1295 TAGATTAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1354
 Db 782 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 727
 Qy 1355 ATGAAAGTTTGAATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1414
 Db 726 ATTATTAATAAATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 667
 Qy 1415 TACAGACGCAATGATATTTATTCATGACCTAAACTGAAATGATTTAAAAATAC 1474
 Db 666 AAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 607
 Qy 1475 CGTTCTTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1534
 Db 606 TTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 547
 Qy 1535 ATCAAGATGATTAAGATTTAGATATTTACATATGAGAAATTAACCATTTGATTA 1594
 Db 546 AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 487
 Qy 1595 TAAATTAAGCATTTATTTGATTTGCAATTAACAAAGATTTAA 1640
 Db 486 AAAAAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 441
 RESULT 7
 CNS03LUB
 LOCUS
 DEFINITION
 829 bp DNA linear GSS 17-MAY-2000
 Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
 036f11 of library G from Tetraodon nigroviridis, genomic survey

[illegible][illegible]

QY	1575	AGAAATACCAAT	1588	
Db	1070	AAWAAHAWMACWT	1083	
RESULT 10				
CNS00396/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
Source				
BASE COUNT				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 128;				
QY	943	AGAAAAATTTTGCAGCGAAAAAATATTTTGGAAATTTTAAAGCGATCTTGGTACCGC	1002	
Db	1099	ARRGGDDTWDRTDRKDDWMTKMTWDRADDRRMGADADRRMAMDDGATWMTATMWW	1040	
QY	1003	ACTTTTGGCAATTTAAACCTGACTCTTTATTAAGTAAGTAATATATCCGTAGATT	1062	
Db	1039	WWWWATDWTWDDKMMWWATAAKTDTJATWMTATWTRAPADGADRGKGRDRDAADADAG	980	
QY	1063	ATAAAGTATGTTAAAAAGCATAAACAATACCTATATATTTAATTCGATATATTT	1122	
Db	979	RRDGGRRRKKDKDDGDDDKGKKKAAAAAWATKWWDDWMDKWKWDGAKDRAD	920	
QY	1123	TGACAGTATTTATTTAATATATTAAGATATATCTATTAAGTAAATATATACATAAAAA	1182	
Db	919	DDGAGGDKDDGKKKADDDDTDFGRKDDDDKXKMDWAKAKTGWGDATWAAWATDMMWGW	860	
QY	1183	AGAGCTAAATATATGATTTGTG-TATTTAAAAAGCATTAAGAAATGAAATGAGACATTA	1241	

Db	B859	ADADMMWTTDAADAADDMMNADRFDAMWKMWDMAWAMPATRADRRDMDGRGKRGARRRDR	800
Qy	1242	TAAAAAAGACGGGTATATCAAACTCTTCTTACACTACTTGATATCTTGATATGATTA	1307
Db	799	KRAADKRPAADRDDAATTWTWTWTTRTDIDDMWKTDITWTWRMADRWDHDDDDDRBA	740
Qy	1302	AGCATTTAATGCATGATGATGGCGATATTTATAACTCAAAATTTATCCATTTGAGAAG	1361
Db	739	GTAARKRKRRTTKRKMRKRDFTWDDADADDTIARDRRRGDDGADAGKKTKGRKRRDR	680
Qy	1362	TTTTGATCTCTTGATGTTGAGCACAATTTTCGGCTGGCTTTCAATTAATATACAGAG	1421
Db	679	ATWRTDPAMWDAAMWTTTDTDTDDMDCRDRRRRGARRRRRTTARAADMWMTWKAMDPAK	620
Qy	1422	ACGAATGAATTTATTCATCAGTCAGCTAAACGAAATGATTTTAAAAATTCOTGCTCTT	1481
Db	619	WMKTRADRMWADPTWTDKAKDRDKRAMRARRRARARAFADRWTKGKTITAT	560
Qy	1482	TATTCATTCGATCAATTTGGAGAAGATTTAATTCAGATGAGATGATACATATACATCA	1541
Db	559	WTFWAARPAWMAWMAWTTATTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT	500
Qy	1542	AATGATTTAAGGATTTAGTATATTTACATATGCGAATTAACCATATGAATTAATTAATA	1601
Db	499	AWTAAAAAAMAAAAAATTTTTTTTTTTTTTTTWWMTAAWMTAWTWTWTWTWTWTWT	440
Qy	1602	AGAGAAATATTTGATTCGAAATTAACAAGATTAATAAGAAGTACAGCTCT	1653
Db	439	TTWTWTWTWMAATTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT	388
RESULT 11			
BH183498/c			
LOCUS	BH183498		
DEFINITION	023 L.O7-rev SmbAc1 Schistosoma mansoni genomic clone 023L07 5',		
VERSION	BH183498		
KEYWORDS	BH183498.1 GI:16288814		
ORGANISM	GSS.		
SOURCE	Schistosoma mansoni.		
REFERENCE	Schistosoma mansoni		
AUTHORS	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;		
TITLE	Strigoidia; Schistosomatidae; Schistosomatidae; Schistosoma.		
JOURNAL MEDLINE	1 (bases 1 to 660)		
COMMENT	Le Paslier,M.-C., Pierce,R.J., Merlin,F., Hirai,H., Wu,W., Williams,D.L., Johnston,D., Loverde,P.T. and Le Paslier,D. Construction and characterization of a Schistosoma mansoni bacterial artificial chromosome library Genomics 65 (2), 87-94 (2000) 20247247 Other-GSSS: 023_L_07-21 Contact: Pierre RJ INSERM U 167 Institut Pasteur de Lille 1 rue du Professeur A. Calmette, 59019-Lille, France Tel: (33) (0)3 20877783 Fax: (33) (0)3 20877888 Email: Raymond.Pierce@pasteur-lille.fr CNS sequencing ID-DG0AA023CF04BP1 Plate: 023 row: L column: 07 Seq primer: M13 reverse primer Class: BAC ends High quality sequence stop: 660. Location/Qualifiers I..660 /organism="Schistosoma mansoni" /strain="Puerto-Rican" /db_xref="taxon:6183" /clone="023L07" /clone_lfp="SmbAc1" /sex="mixed" /dev_stage="cercariae" /lab_host="Biomphalaria glabrata"		
FEATURES			
Source			

/note="Vector: pBel0BAC 11; Site 1: Hind III; Partially Hind III digested and size-selected S. mansoni cercarial DNA was ligated into Hind III digested pBel0BAC 11 vector and used to transform E. coli DH10B. The complete library contains 23808 clones from 4 independent sizing-ligation transformations. Average insert size ranges from 70-127 kb and genome coverage is 7.9-fold."

Query Match 5.0%; Score 93; DB 17; Length 660;
Best Local Similarity 43.6%; Pred. No 3e-06;
Matches 278; Conservative 35; Mismatches 323; Indels 1; Gaps 1;

QY	1061	TTATAAAGTATGTTAAAAACGATGAAACAACTATATATTTAATCTGATATATA	1120
DB	640	TTTTTAAAAAATAAAAAAAAAAAAAAAAAAATTTTAAATATTTTAAATATTT	581
QY	1121	TTTGACAGTGTATTTATATATTAAGAGTATATCTATTAGCTTAAATTTAACTAATA	1180
DB	580	TATTTTAAWAAAWTAAAAAATAAAAAAAAAAATAATTTTATTTTAAAAA	521
QY	1181	AAGAGGATATATATGATGTGCTATTTAAAAAGCAATTTGAATAAATAATAGACTT	1240
DB	520	AATADTKAAWTTTAAAAAATAAATATATWAAAAAATTAATTTTAAATTTT	461
QY	1241	ATAAAAAAGCGGTGATATCAAAATCTTCTTACATATACTGCATTACTTTGATATAGATA	1300
DB	460	TWAAAAATTTTAAWAAAWAAATATATATTTAAATATATTTAAAAAATAATATATTA	401
QY	1301	AAGCATTAATATGTCATGATGTCGCGATTTATTAACCTCAATTTATCATTTGATGATA	1360
DB	400	AGGCGTATATAAAAAGCTWAAARAAAAAGTAAAAAATAATATTAATAAATGAAAAA	341
QY	1361	GTTTGTGATCTCTTGATGCTGTGACACATTTTCGGCTGGCGTTCAATTAATAATATATACGA	1420
DB	340	TATATTAATWACAC-CGTACTCTKTKTDARAAAAAGTTTCCCAAGAAAAAATAA	282
QY	1421	GACCAAAATGATATTTATCATCACCTAAACCTGAATAAGATTTAAAAAATACTCGTCT	1480
DB	281	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	222
QY	1481	TTATTCATTCATCAATTTGGAGAGATTTAATTCGATGAGATGATCAATATACATCAAG	1540
DB	221	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	162
QY	1541	AATGATTAAGAGATTCATATTAATTAACATATGGAATAATCACTATGATATATATAAT	1600
DB	161	AAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	102
QY	1601	AAGAGATATATTTGATTCGAATATACAAAAAGATATTAAGAAGATCGCTTTAAAT	1660
DB	101	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	42
QY	1661	ATGTCATTCATGATGATGTTTAAAAAGTAGATGAATA 1697	
DB	41	AAAAATAAAAAATTTAATATWAAAAATTTAAAAAATAA 5	

RESULT 12
CNS070NJ/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

660 bp DNA linear GSS 10-OCT-2001
T3 end of clone 023CF04 of library SBB41 from strain Puerto-Rican
of Schistosoma mansoni, genomic survey sequence.
AL620449
AL620449.1 GI:16034591
GSS.
Schistosoma mansoni.
Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigoida; Schistosomatidae; Schistosomatidae; Schistosoma.
1 (bases 1 to 660)
Le Pastier,M.C., Pierce,R.J., Merlin,F., Hirai,H., Wu,W.,

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	COMMENT	FEATURES	BASE COUNT	ORIGIN
Williams, D.L., Johnston, D., Loverde, P.T., and Le Paslier, D.									
Construction and characterization of a Schistosoma mansoni									
Genomics 65 (2), 87-94 (2000)									
20247247									
10783255									
2 (bases 1 to 660)									
Genoscope.									
Submitted Submission									
Direct (05-OCT-2001) Genoscope - Centre National de Sequencage :									
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr									
- Web : www.genoscope.cns.fr)									
Partially Hind III digested and size-selected S. mansoni cercarial									
DNA was ligated into Hind III digested pBeloBAC 11 vector and used									
to transform E. coli DH10B. The complete library contains 23808									
clones from 4 independent sling-ligation transformations. Average									
insert size ranges from 70-127 kb and genome coverage is 7.9-fold.									
Location/Qualifiers									
1..660									
/organism="Schistosoma mansoni"									
/strain="Puerto-Rican"									
/db_xref="taxon:6183"									
/clone="023CF04"									
/clone.lib="SmbAC1"									
/note="end : T3"									
BASE COUNT	139 a	16 c	9 g	446 t	50 others				
ORIGIN									
Query Match	5.0%: Score 93; DB 17; Length 660;								
Best Local Similarity	43.6%: Pred. No. 3e-06;								
Matches	278: Conservative	35: Mismatches	323: Indels	1: Gaps	1: Gaps				
1061	TTATAAAGTATGTTAAAAACGAGTAAACAACTATATATATTTAATTCGATTATA	1120							
11111	11111	11111	11111	11111	11111				
640	TTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAATTTTATMTATTTTATATTTATT	581							
1121	TTTGACAGTGATTTATTTATATATTTAGACATATATCTATTAGCTTAAATTAACATAAA	1180							
580	TATTTTAAWAAWMTAAAAAATAAAAAAAAAAAAAAAAAATTTTATTTTAAAAA	521							
1181	AAAGAGTAAATATATGATGTTGCTATTTAAAAAGCATTTGAATAATGAATAGACACTT	1240							
520	AAMADTKAAWTTTAAAAAATAAATAATATATWATAAAWATATATATTTTATTTT	461							
1241	ATAAAAAAGACGATGATCAAACTCTTCTTCAATCAATCTACTGATTTGATATAGATA	1300							
460	TWAAAAAATTTTAAWAAWAAATATATATWATAATATTTAAAAAATAATATATATATA	401							
1301	AAGCATTAATATGATGATGATGCGCATATTAATTAACATCAATTTATCCATGATGAAA	1360							
400	AGGCTWTTAAAAAAWMTWAAARAAAAAGTAAAAAAATATTAATAAAATGAAAAWMTWAAA	341							
1361	GTTTTCATCTCTTGATGATGTTGACGACAATTTGCGCTGGCTTTCATTAATAATATACGA	1420							
340	TATATMATATCAC-GCTACTCTKTKIDAKRAAAAGATTTCCCAAGAAAAAATAAAAAA	282							
1421	GACCAAAATGATATTTTATCATCAGCTAAATATGAAAAATGATTTTAAAAAATCTGTTCT	1480							
281	AAAAAATAA	222							
1481	TTATTCATTCGATCAATTTGGAAGATTTTAATTAATGATGAGATGATGATCAATCATCAAG	1540							
221	AAAAAATAA	162							
1541	AAATGATTAAGGATTTAGATTAATTAACATATGAGAAATATACATATGATTAATAAT	1600							
161	AAGAAAAAATAAA	102							
1601	AAGAGATATATTTGATTTTGAATATACAAAAAGATATTAAGAAAGATGACGCTTAAAT	1660							
101	AAAAAATAAAAAAAAAATAATTAATAAAAAAAAAAAAAAAAAATAATTAATTAATAATTAAGAAAAA	42							
1661	ATGTCATTTGATGATGTTTAAAAAGTATGATGAAAA	1697							


```
Db 1046 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1105
QY 1279 TTGCATTACTTGATATAGATTAAGCATTAAATGGTGAATGGCATATTAATAAC 1338
Db 1106 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1165
QY 1339 TCAATTATCCATGATGAAGTTGTCCTTCTTGATGTTGACACAAATTCGGCTGG 1398
Db 1166 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1225
QY 1399 GCTTCATTAATAATATCAGAGCAATGAATATTATCATCTACGCTAAACCTGAAT 1458
Db 1226 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1285
QY 1459 GATTTTAAAAAATCTCTCTTTTATTCATTCGATCAATGGAGAAATTTAATTACGAT 1518
Db 1286 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1345
QY 1519 GAGATGAGTACATACATCAGAAATGATTAAGATTAGATTAATTACATATGGAGAA 1578
Db 1346 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1405
QY 1579 ATAAACCATATGAATATAATAAAGAAATATATGATTGGAATAACAAAGATATAA 1638
Db 1406 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1465
QY 1639 AAGAAAGTCAGCTCTTAAAAATATCTGCATTGATGATGTTTAAAGTAGATGAAAAAT 1698
Db 1466 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1525
QY 1699 TTATTGATGAGAGGATTTGCACTAAAGATATGAAAAATATCGTATGAAAAATCTATG 1758
Db 1526 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1585
QY 1759 ATGATCCAGATGATGCAATTAAGAAATCACAATTCGACAGAAATGCTATGCTTCC 1818
Db 1586 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1645
QY 1819 GCATTAAAAAACAAGCAAAAAAGAGA 1844
Db 1646 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1671
```

Search completed: July 17, 2003, 06:38:58
Job time : 2708.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 21:08:10 : Search time 377.725 Seconds
(without alignments)
9896.940 Million cell updates/sec

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAGCAAAAT.....AAGACTCGAGCAGATCT 1660

Scoring table:
IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1657	99.8	1660	20	AAV74275
2	109	6.6	6292	22	AA546735
3	94.8	5.7	778	24	ABO15588
4	94.8	5.7	778	24	ABO15589
5	93.6	5.6	6644	20	AA33181
6	93.6	5.6	7372	20	AA33182
7	93.6	5.6	7797	20	AA33180
8	93.6	5.6	7996	20	AA33184
9	90.2	5.4	14006	24	ABL33958

C 10	89.8	5.4	6668	24	ABL33697	Human immune syste
C 11	88.2	5.3	975	24	ABO29508	Oligonucleotide fo
C 12	88.2	5.3	975	24	ABO29509	Oligonucleotide fo
C 13	87.6	5.3	7442	22	AA546686	Tumour suppressor
C 14	86.6	5.2	15548	24	ABL34155	Human immune syste
C 15	83.8	5.0	9539	22	AA545347	Chemically pretrea
C 16	83.8	5.0	9539	24	ABK28180	DNA transcription
C 17	83.6	5.0	7571	24	ABL32527	Human immune syste
C 18	83.6	5.0	7823	22	AA545490	Chemically pretrea
C 19	83.6	5.0	7823	24	ABL34061	Human immune syste
C 20	83.6	5.0	7823	24	ABK28418	DNA transcription
C 21	83.6	5.0	7823	24	ABK31493	Signal transductio
C 22	82	4.9	7814	22	AA546530	Tumour suppressor
C 23	81.6	4.9	11745	24	ABK28332	DNA transcription
C 24	80	4.8	7597	24	ABL33013	Human immune syste
C 25	80	4.8	12237	24	ABL34358	Human immune syste
C 26	79.8	4.8	5276	24	ABL22150	Human immune syste
C 27	79	4.8	6419	24	ABL32267	Human immune syste
C 28	78.4	4.7	2943	21	AA470229	Plasmodium falcipa
C 29	78.4	4.7	15674	24	ABL70514	Chemically treated
C 30	78.4	4.7	15674	24	ABL32363	Human immune syste
C 31	78.4	4.7	15674	24	ABL34477	Human immune syste
C 32	78.2	4.7	9964	24	ABL32098	Human immune syste
C 33	76.8	4.6	6314	24	ABL54314	Chemically treated
C 34	76.6	4.6	5641	24	ABL33396	Human immune syste
C 35	76.4	4.6	1200	24	ABO39210	Oligonucleotide fo
C 36	76.4	4.6	1200	24	ABO39211	Oligonucleotide fo
C 37	76.2	4.6	8079	24	ABL22313	Chemically treated
C 38	76	4.6	15387	24	ABL32184	Human immune syste
C 39	75.4	4.5	1998	21	AA470212	Plasmodium falcipa
C 40	75.2	4.5	6159	24	ABN80236	Human chemically m
C 41	75.2	4.5	7461	24	ABL33784	Human immune syste
C 42	75	4.5	1030	24	ABO42518	Oligonucleotide fo
C 43	75	4.5	1030	24	ABO42519	Oligonucleotide fo
C 44	74.2	4.5	18997	24	ABK33948	Human DNA for stag
C 45	74.2	4.5	18997	24	ABL32570	Human immune syste

ALIGNMENTS

RESULT 1
ID AAV74275 standard; DNA: 1660 BP.

AAV74275:

16-JUN-1999 (first entry)

Pantoea citrea cryptic plasmid ps.

Cryptic plasmid; ps; higher temperature; growth; elimination; mobilization; ds.

Pantoea citrea.

Key location/Qualifiers
480..1595
/*tag= a
/note= "largest open reading frame"

W09859054-A1.

30-DEC-1998.

22-JUN-1998; 98WO-US12945.

23-JUN-1997; 97US-0876132.

(GENV) GENENOR INT INC.

Causey SC, Fowler T;

[illegible]

RESULT 3
ABQ15588/C
ID ABQ15588 standard; DNA: 778 BP.
XX
AC ABQ15588;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 2179.
XX
KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KM drug; side effect; cancer; central nervous system; cardiovascular;
KM gastrointestinal; respiratory system; single nucleotide polymorphism;
KM SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP10074.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI: 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
XX Sequence 778 BP; 111 A; 24 C; 29 G; 614 T; 0 other;

Query Match 5.7%; Score 94.8; DB 24; Length 778;
Best Local Similarity 46.4%; Pred. No. 4.7e-08;
Matches 345; Conservative 0; Mismatches 397; Indels 2; Gaps 1;

QY 55 CCGACTTAAACCAATGAGTGAAGAGAAAGAAAAATTAATACATATTTTGGAGTAG 114
DB 776 CCGAAAAAATAATACGAAAAAATAATTAATAAATAATATATATAAAAAA 717
QY 115 TAAAAAGAGAAAAAATAAAGAGAGCTCGCTTACAGCTGAAAAACAGAAATTA 174
DB 716 AAATTACGAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTA 657

QY 175 TAAAAAGAGAGAGCTGTGATTTTAAATGAAATCGTAGAAAAAGAAATTTTAAATTT 234
DB 656 TAAAAAATAAATAATTAAGAAAAAATAATATATATAAATAAATAATTAATAAATAA 597
QY 235 CATTTTCAGAGATTAATTTGTTGAAGTGTGATGAAAAATCTAGATTAATAAATGAGATC 294
DB 596 AAAATTCGAAAAAATAAATAATTAATAAATAAATAAATAATTAATAAATAAATAA 537
QY 295 AAAATGTGTAATTTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 354
DB 536 ACGAAAAAATAAATAATTAATAAATAAATAAATAATTAATAAATAAATAAATAATTA 477
QY 355 TATTTTATTTTTCGAGAGATTAATAATCTGAGTGAAGAAATAGTTTCGAGAGCAAAA 414
DB 476 AAAAAAATAATTAAGAAAAAATAAATAATTAAGAAAAAATAAATAATTAAGAAA -AAAA 419
QY 415 AACCTTCGCGTTTTCCTCAATGACTTTGAAAAAATTCATTTGACGCGTAGCGAANC 474
DB 418 AAAATTTCAAAAAACGAAAAAATAATTAATAAATAAATAAATAAATAAATAATTA 359
QY 475 TTTGAATTTTTCATTTGGAATTTGAATAATTAAGCAAAAGAACTCAAAATGAGAAA 534
DB 358 AAAAAAATAAATAATTAAGAAAAAATAAATAATTAATAAATAAATAAATAATTA 299
QY 535 AATATTTATTAATAAAGAGAGATCGATATGATTTTAAAGCAGAAACTGACATTGA 594
DB 298 AAAATTCGAAAAAATAAATAATTAATAAATAAATAAATAAATAAATAATTA 239
QY 595 ATGAAAAAATAAATTTGAAAAAATCTATGCTGAGAGCAATTAATAACCAAAAAATG 654
DB 238 AAAAAAATAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAATTA 179
QY 655 GAATCAACCGGTGTTGTTTACAAATGACGATGAAGAAATGATGAATAATATCAACC 714
DB 178 AAAAAAATAATTAAGAAAAAATAAATAATTAAGAAAAAATAAATAATTAACAAAAA 119
QY 715 TCGATGTTAATGAGAAACAGCAGGTCAATATAGAAATTAATCAAAATAAAGTTGAGC 774
DB 118 AAATTAATTAATAAATAAATAAATAATTAAGAAAAAATAAATAAATAAATAAAT 59
QY 775 ATAGTAATCAGATGATCTAGTAA 798
DB 58 ATAAAAAATAAATAATTAATAA 35

RESULT 4
ABQ15589
ID ABQ15589 standard; DNA: 778 BP.
XX
AC ABQ15589;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 2180.
XX
KM Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KM drug; side effect; cancer; central nervous system; cardiovascular;
KM gastrointestinal; respiratory system; single nucleotide polymorphism;
KM SNP; cell differentiation; ds.
XX
XX Homo sapiens.
XX
OS WO200218632-A2.
XX
PN 07-MAR-2002.
XX
PD 01-SEP-2001; 2001WO-EP10074.
XX
PF 01-SEP-2000; 2000DE-1043826.
XX
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPIG-) EPIGENOMICS AG.

KM Cowpox virus; bsr: viral vector; expression; apoptosis; resistance;
 KM crma; bcl-2; bcl-xl; FLIP; survivin; IAP; ILP; adenovirus; cancer;
 KM autoimmune disease; graft rejection reaction; inflammation;
 KM inflammatory disease; ss.

OS Synthetic.

OS Homo sapiens.

PN WO9913073-A2.

PD 18-MAR-1999.

XX 07-SEP-1998; 98WO-JP04010.

XX 08-SEP-1997; 97JP-0259235.

XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.

XX Hamada H;

DR WPI; 1999-243728/20.

XX New apoptosis-resistant virus-sensitive cell

PS Example 3; Page 46-49; 51pp; English.

CC The present invention describes an apoptosis-resistant virus-sensitive
 CC cell line into which an apoptosis resistance gene has been introduced.
 CC The recombinant viruses generated are capable of expressing apoptosis-
 CC associated genes. These can then be used in a variety of diseases for
 CC which the induction of apoptosis is therapeutic. The recombinant viruses
 CC are useful as vectors for gene therapy which can be applied to cancer
 CC therapy for destroying cancer cells selectively, the treatment of
 CC autoimmune diseases and graft rejection reaction, and apoptosis induction
 CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
 CC encountered the problem where if an adenovirus vector capable of
 CC expressing an apoptosis-associated gene is introduced into animal cells,
 CC the cells producing the virus will be destroyed because the period of
 CC time required to induce cell death by apoptosis is shorter than that
 CC required to replicate and produce the virus, resulting in failure to
 CC obtain a recombinant virus having the integrated apoptosis-associated
 CC gene. In this invention an apoptosis-resistant 293 cell line (having an
 CC apoptosis resistant gene introduced) is established and overcomes the
 CC problem. The present sequence represents the base sequence of the
 CC plasmid pR-Bcl 2-1-hCD 25, which contains the human Bcl-2 gene, and
 CC is used in an example from the present invention.

CC Sequence 7996 BP; 2463 A; 2015 C; 1829 G; 1689 T; 0 other:

Query Match 5.6%; Score 93.6; DB 20; Length 7996;

Best Local Similarity 45.9%; Pred. No. 9.4e-08;

Matches 321; Conservative 0; Mismatches 379; Indels 0; Gaps 0;

OY 7 ACACAGCGCAATTTGAAAAATAGATTAATTTTGCAGCTTTTAAAGCCGCTTAAC 66
 DB 5139 AA 5198
 OY 67 AAATGAGTGAAGAGAAAGAAATTAATATCATTTTGTAGTAAAGAGAAAG 126
 DB 5199 AA 5258
 OY 127 AAAAAATTAAGACGCTCGGCTTACAGCTGAAAAACCAAAATAATTAAGAAAGA 186
 DB 5259 AA 5318
 OY 187 GACTGTGATTTTATGAAATCGTGAGAAAGAAATTTTAAATTTTCATTTTCGAGG 246
 DB 5319 AA 5378
 OY 247 ATTAATTTGTTGTAAGTGAATAATCTACATTAATAATTCAGATCAAAATGCTTG 306
 DB 5379 AA 5438

OY 307 AATTTGACATTAATGAAATACCTAGTAAATCAATTAATGAGGCTTGTCTATTTTATTTG 366
 DB 5439 AA 5498

OY 367 CGAAGATTGAAATATCTGAGTGAAGAAATTAATTTGCGAGACCAAAACCTTGCCGT 426
 DB 5499 AA 5558

OY 427 TTTTTCAAATGCTTTGCAAAATTCATGTGTGAGCGGTAGCAAACTTGAAATTTT 486
 DB 5559 AA 5618

OY 487 TACATTTGAAATTTGAAATTAAGCAAAAGAACTCAATGAAATTAATTAATTA 546
 DB 5619 AA 5678

OY 547 AAAAAAGCAGATCGATATGATTTTAAAGCAGAAACTGCAATTTGAAATGAAGAAAG 606
 DB 5679 AA 5738

OY 607 ATTTGAAATAATCTATGCTGAGATGAATTAAGCAAAATTTGGAACCTCAACCG 666
 DB 5739 AA 5798

OY 667 GTGTGTTTTCAGAAATGACGATGAAGAAATGATGAAGAA 706
 DB 5799 GAGATTTTCAGAGAGCTAAGGAAGCTAAATGAGAAAGAA 5838

OY 667 GTGTGTTTTCAGAAATGACGATGAAGAAATGATGAAGAA 706
 DB 5799 GAGATTTTCAGAGAGCTAAGGAAGCTAAGGAAGCTAAGGAAGAA 5838

RESULT 9
 ABL33958/c
 ID ABL33958 standard; DNA; 14006 BP.

XX ABL33958;
 DE 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 1931.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antidiabetic; anti-HIV; anticonvulsant; ophthalmologic;
 KW antirheumatic; antidiabetic; antidiabetic; antidiabetic;
 KW antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 KW acute myeloid leukaemia; Alzheimer's disease; arteriosclerosis; anemia;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.

OS Homo sapiens.

PN WO200200928-A2.

PD 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation

PS Claim 1; SEQ ID NO 1931; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated

CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/nutritional bowel
 CC diseases. The present sequence is a gene of the invention.

XX Sequence 14006 BP; 3278 A; 155 C; 3257 G; 7313 T; 3 other;

Query Match 5.4%; Score 90.2; DB 24; Length 14006;

Best Local Similarity 46.2%; Pred. No. 4e-07;

Matches 299; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

```

QY 4 TCTACACAGCAAAATGAAAAATAGATAATTTTGGCAGGTATTAAAGCCGACTAA 63
DB 2562 TATTAACATATTAATCTTAACAAACCTTCCTTAATTAATTTCTCCTCTAC 2503
QY 64 AACAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 123
DB 2502 TATTAACATATTAATCTTAACAAACCTTCCTTAATTAATTTCTCCTCTAC 2443
QY 124 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 183
DB 2442 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2383
QY 184 AGAGAGCTGATTTTAAATGGAATCGTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
DB 2382 AACCAATATTAATCTTAACAAACCTTCCTTAATTAATTTCTCCTCTAC 2323
QY 244 GGGATTAATTTGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 303
DB 2332 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2263
QY 304 TTGAATTTGACATTTGAAATAGCTAGTATTCATTAATGGGGTGTCTATTATT 363
DB 2262 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2203
QY 364 TTGCGAATGGAATTCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
DB 2202 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2143
QY 424 CGTTTTCATTAATCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
DB 2142 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2083
QY 484 TTTTACATTTGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
DB 2082 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2023
QY 544 AATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
DB 2022 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1963
QY 604 AAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 650
DB 1962 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916

```

RESULT 10

ABL33697/C

ID ABL33697 standard; DNA; 6668 BP.

XX ABL33697;

XX 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1670.

XX Human; immune system disease; cytosine methylation; antistimatic;

KW antileukemic; antileukemic; antileukemic; antileukemic;

KW antileukemic; antileukemic; antileukemic; antileukemic;

KW antileukemic; antileukemic; antileukemic; antileukemic;

KW antileukemic; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.

OS Homo sapiens.

PN WO20020928-A2.

PD 03-JAN-2002.

PF 02-JUL-2001; 2001WO-EP07537.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIDEMIOLOGICS AG.

PI Olek A, Plepenbrock C, Berlin K;

DR WPI: 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation

PS Claim 1; SEQ ID NO 1670; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/nutritional bowel
 CC diseases. The present sequence is a gene of the invention.

XX Sequence 6668 BP; 1329 A; 328 C; 1736 G; 3274 T; 1 other;

Query Match 5.4%; Score 89.8; DB 24; Length 6668;

Best Local Similarity 44.7%; Pred. No. 4.4e-07;

Matches 349; Conservative 0; Mismatches 432; Indels 0; Gaps 0;

```

QY 7 ACACAGCAAAATGAAAAATAGATAATTTTGGCAGGTATTAAAGCCGACTTAAAC 66
DB 3352 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3293
QY 67 AATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 126
DB 3292 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3233
QY 127 AAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186
DB 3232 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3173
QY 187 GAGCTGATTTTAAATGGAATCGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
DB 3172 AAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3113
QY 247 AATTAATTTGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306
DB 3112 AAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3053
QY 307 AATTTGACATTTTGAATAGCTAGTATTCATTAATGGGGTGTCTATTATTATT 366
DB 3052 AAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2993
QY 367 CGAGATTTGAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426
DB 2992 AAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2933
QY 427 TTTTTCATTAATCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486

```


RESULT 12

ID	AHQ29509	standard; DNA; 975 BP.
XX	ABQ29509;	
DT	12-JUL-2002	(first entry)
DE	Oligonucleotide for detecting cytosine methylation SEQ ID NO 16100.	
KW	Human: cytosine methylation; 5'-Cpg-3'; uracil; cytosine; diagnosis; drug; side effect; cancer; central nervous system; cardiovascular; gastrointestinal; respiratory system; single nucleotide polymorphism. SNP; cell differentiation; ds.	
OS	Homo sapiens.	
PN	WO200218632-A2.	
PD	07-MAR-2002.	
PF	01-SEP-2001; 2001MO-EPI0074.	
PR	01-SEP-2000; 2000DE-1043826.	
PR	05-SEP-2000; 2000DE-1044543.	
PA	(EPIG-) EPIGENOMICS AG.	
PL	Olek A, Piepenbrock C, Berlin K, Guetig D;	
DR	WIJ; 2002-371829/40.	
PT	Determining the degree of cytosine methylation in genomic DNA, useful for diagnosis and prognosis, comprises selective hybridization of amplicons from chemically treated DNA -	
XX	Claim 12; 56pp + Sequence Listing; 56pp; German.	

This invention describes a novel method for determining the degree of methylation of a particular cytosine in a motif 5'-Cpg-3', present in a genomic sample of DNA. The sample is treated chemically to convert cytosine (C) but not methylated C, to uracil, then part of the genomic DNA that contains the target C is amplified to form a labeled amplicon. The amplicon is hybridised to two classes, each with at least one member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the degree of hybridisation to both classes is determined from the label on the amplicon. From the ratio of labels hybridised to the two classes of oligomers, the degree of methylation is calculated. The method is used: (1) for diagnosis and/or prognosis of side effects of therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders of the central nervous, cardiovascular, gastrointestinal and respiratory systems etc., particularly by detecting mutations or single nucleotide polymorphisms (SNPs); and (ii) for differentiation of cell or tissue types and for investigating cell differentiation. The method allows the methylation status of many C residues to be determined simultaneously. AB013410-AB041121 represent genomic DNA sequences used to illustrate the method for determining the degree of cytosine methylation described in the disclosure of the invention.

```
Sequence          975 BP;   780 A;    51 C;   32 G;  112 T;    0 other;

Query Match              5.3%; Score 88.2; DB 24; Length 975;
Best Local Similarity   45.9%; Pred. No. 7, Le-or;
Matches      379; Conservative     0; Mismatches 438; Indels       8; Gaps.    2
```

YY	22 AAAAATAGATAAAATTTCGCAGGTATTAACCGCACCTAAAACAATGACTGAAGAAG 81
DB	126 AAAAAATTAATAAACCAACAACATTTAAAAATTTAAAAACCAACAACGAATAAAA 185
YY	82 AAAGCAAATAATCATATTTTGAGTGTGTAAGAGAAAGAAAGAAATTAAGAAG 141

Db	186	AAATRAAAAAATRAAAAAACGATTAATAATTAATAAATCAACGATTAATAATA	245
Oy	142	ACCTCGGCTTAACAGTCGAAAAACGAAATTAATAAAGAAGAGACTGATTTTAA	201
Db	246	AAAAATTAATAACAAACGATTAATAATAATAATAATAAAAAACGATTAATAATA	305
Oy	202	TGCAAAATCGAGAGAAAAAATTTTAAATTTTCATTTTCAGAGGATTAATTTGTAA	261
Db	306	ATAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATA	365
Oy	262	GTTGATGAAAAATCTGATTAATAAATTCGATCAAAATTAATGTTGTAATTCATTTTG	321
Db	366	AAACAAACGATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATA	425
Oy	322	AAATAGCTAGTATCAATTAATGGGGTTTGTCTATTTTATTTTTCGAAAGTTGAAATC	381
Db	426	AAACGATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA	485
Oy	382	TGAGTGAAGAAATATGTTGGCAGACAAAAAACCTTGCCGTTTTCATTAATGACT	441
Db	486	GAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATA	545
Oy	442	TTGGAAAAAATTCATGTTGAGGGGTAGCAAACTTGAAATTTTTCATTCGAAATTTG	501
Db	546	AAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA	602
Oy	502	AAAAAATTAAGCGAAAAAAGCACTCAATGAAAAAATTAATTAATAAAGAAGATCGG	561
Db	603	AAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATA	662
Oy	562	ATATGATTTTAAAGCAGAAACTGACATTTGAATGAATAAATAAAGATTGGAAAAATCT	621
Db	663	ATAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA	722
Oy	622	ATGCTGAGATGATTAATAAGCAAAAAATTTGGCACTCAACCCGGTGTGTTTAAATA	681
Db	723	AAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATA	782
Oy	742	AAATATGGAATTAATTCAAAAATTAAGTTGAGCATAGTAATTCAGATGATCTAGTAACG	801
Db	838	ATAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA	897
Oy	802	GACTATTAGCTGAGCACTCGAAATTAATTTGATTAACACAGAA	846
Db	898	AAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA	942
RESULT 13			
DD	AA546686/c	AA546686 standard; DNA; 7442 BP.	
XX	AA546686;		
XX	AC		
XX	18-DEC-2001	(first entry)	
XX	DT		
XX	TT	Tumour suppressor gene derived chemically modified sequence #409.	
XX	DE		
XX	XX	Human; tumour suppressor gene; oncogene; antitumour; cytostatic;	
XX	KW	cancer; tumour; Cpg dinucleotide; single-nucleotide polymorphism; SNP;	
XX	cytosine methylation; ds.		
XX	OS		
XX	Homo sapiens.		
XX	PN	W0200168912-A2.	
XX	PD	20-SEP-2001.	
XX	PF	15-MAR-2001.2001WO-EP02955.	

CC of HIV infection, neurodegenerative disorders, graft-versus-host disease
CC aging, glomerular disease, Lewy body disease, arthritis,
CC arteriosclerosis, solid tumours and cancers.

Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 other;

Query Match	5.0%;	Score	83.8;	DB	22;	Length	9539;
Best Local Similarity	43.0%;	Pred. No.	5.3e-06;				
Matches	463;	Conservative	0;	Mismatches	612;	Indels	2;
						Gaps	1;

QY	4	TCCTCACAGGCAATTTGAAAAATAGTAAATTTTCGCGAGTATTAAAGCCGCTTAA	63
Db	1116	TCCTCTAAAAAAGAAAAAAGAAAAACGAAAAACGAAAAATATAAAAAATAAAAA	1057
QY	64	ACCAATGAGTGAAGAAAGAAAAAATAATTAATACATATTTTGAGTTAGTAAAGAGA	123
Db	1056	AAAAAAGAAAAAATATAAAAAAGAAAAAATATTAACAAAAAAGAAAAAAGAAA	997
QY	124	AAGAAAAAATTAAGAAAGACCTCGGCTTACAGTCGAAAAACGAAATTAATAAAGAA	183
Db	996	TAAAAAAATTTAAAAAAGAAAAAATTAATTAATTAAGAAAAAATTAATTAAT	937
QY	184	AGAGACGTGTATTTTATGGAATCGAGAGAAAGAAATTTTAATTTTCATTTTGA	243
Db	936	AAACAAAAAAGAAAAAAGCGAAAAATTAATCTACACCTTAATATCCCAATCTTTAAAA	877
QY	244	GGGATTAATTTGTGTAGTTGATGAATAATCTAGATTAATAAATGACAGATCAAAATGTG	303
Db	876	ACTTAATTAATAAATCAGCTTAAATACCAAAATTTCAAAACCACTTAACCAACAAATCA	817
QY	304	TTGAAATTTGACATTTTGAATTAAGTATATCAATTAATGGGGTTTGCTATTTTAT	363
Db	816	TACCCCATTTCTAAAAAAGAAAAAAGCTAACCAAACTTAATACACACACTATAA	757
QY	364	TTGGAGATTTGAATACTGATGAAGAAATAGTTGGCAGAGCAAAAAACCTTCGC	423
Db	756	TCCCACTACTCAAAAAAATTAACAAAAAATCACTACGAAAAACATATTAACAATA	697
QY	424	CGTTTTTTCAAAATGACTTTGGAAAAAATTCATTGTGACGCGTACGAACCTTTGAATTT	483
Db	696	AACGAAATGTACACACTACACTCCAACTAACACAAAAAAGCCCTACTCAAAAAA	637
QY	484	TTTTTACATTGAAATTTGAAAAAATTAAGCAAAAGAACTCAATTGAAAAATTTAT	543
Db	636	AAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	577
QY	544	ATTAATAAAGAGATCGGATATGATTTTAAACAGAAAACTGACATGATGAAAAA	603
Db	576	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	517
QY	604	AAGATTTGAAAAAATCTATGCTGAGAGTGAAATTAAGCAAAAAAATTTGGAACTCAAC	663
Db	516	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	457
QY	664	CCGCTGTTTTRGAATGACGATGAAGAAATGATGAATAATATCAACTCGATGTTA	723
Db	456	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	397
QY	724	ATGAGAAAGACGAGTCAATATGGAATTAATCAAAAAAATTAAGTTGAGATAGTAAT	783
Db	396	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	337
QY	784	CAGATGATTAAGTAACGAGACTATTAGAGTGAGACTCGAAATAGTTTGTATTAACAA	843
Db	336	AAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	277
QY	844	GAAAGTCCCTTCTGTTTGTATTTGTGAGAGAAATTCAGCACTGAGAAAAAGAACTGATA	903
Db	276	CAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	217
QY	904	ATGCAAGAAAGATTAAGATTTTGCATATCAATGAAGCAAAAAAATTAAGAGCTTTGAT	963
Db	216	ATACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGCT	157

[illegible]

Search completed: July 17, 2003, 05:07:52
Job time : 392.725 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2003, 02:22:20 : Search time 680.973 Seconds
(without alignments)
5635.000 Million cell updates/sec

Title: US-08-876-132-2
Perfect score: 1847
Sequence: 1 AGATCTCAACAGCAGTTTAAAAA.....AACAAGCAAAAGAGATCT 1847

Scoring table: IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 1448851 seqs, 1038787357 residues

Total number of hits satisfying chosen parameters: 2897702

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq1.*
11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq3.*
13: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.6	4.4	5689	US-10-239-676-90	Sequence 90, Appl
2	77.6	4.2	11812	US-10-239-676-210	Sequence 210, Appl
3	74.6	4.0	9539	US-10-239-676-51	Sequence 51, Appl
4	74.4	4.0	755	US-10-027-632-128109	Sequence 128109,
5	74.4	4.0	9515	US-10-239-676-160	Sequence 160, Appl
6	73.2	4.0	8085	US-10-172-086-40	Sequence 40, Appl
7	72.8	3.9	3252	US-10-027-632-113786	Sequence 113786,
8	72.8	3.9	3252	US-10-027-632-113787	Sequence 113787,
9	72.8	3.9	3252	US-10-027-632-113788	Sequence 113788,
10	72.6	3.9	15732	US-10-239-676-95	Sequence 95, Appl
11	72.6	3.9	15732	US-10-239-676-96	Sequence 96, Appl
12	72.6	3.9	15732	US-10-239-676-96	Sequence 96, Appl
13	72.6	3.9	15732	US-10-239-676-96	Sequence 96, Appl
14	69	3.7	8801	US-10-239-676-143	Sequence 143, Appl
15	67.8	3.7	5979	US-10-239-676-18	Sequence 18, Appl
16	67	3.6	640681	US-09-790-988-1	Sequence 1, Appl

17	66.4	3.6	335913	US-09-754-853A-2	Sequence 2, Appl
18	66.4	3.6	335913	US-09-754-853A-3	Sequence 3, Appl
19	66.2	3.6	7195	US-10-239-676-30	Sequence 30, Appl
20	66	3.6	53332	US-09-801-861-3	Sequence 3, Appl
21	66	3.6	53332	US-10-224-562-3	Sequence 2, Appl
22	66	3.6	335913	US-09-754-853A-2	Sequence 3, Appl
23	66	3.6	335913	US-09-754-853A-3	Sequence 3, Appl
24	65.6	3.6	960	US-10-198-846-6381	Sequence 6381, Ap
25	65.6	3.6	11047	US-10-195-144-87	Sequence 87, Appl
26	64.8	3.5	11047	US-10-239-676-188	Sequence 188, App
27	64.8	3.5	11836	US-10-239-676-102	Sequence 102, App
28	64.6	3.5	12968	US-10-239-676-202	Sequence 202, App
29	64.4	3.5	516	US-09-960-352-5785	Sequence 5785, Ap
30	64.4	3.5	11836	US-10-239-676-101	Sequence 101, App
31	64.2	3.5	17421	US-10-239-676-54	Sequence 54, Appl
32	64	3.5	1431	US-09-774-414-2	Sequence 2, Appl
33	63.8	3.5	2086	US-10-027-632-250719	Sequence 250719,
34	63.8	3.5	2086	US-10-027-632-250720	Sequence 250720,
35	63.8	3.5	5979	US-10-239-676-18	Sequence 18, Appl
36	63.6	3.4	431	US-09-960-352-5558	Sequence 5558, Ap
37	63.4	3.4	628	US-10-027-632-113152	Sequence 113152,
38	63.4	3.4	6665	US-10-239-676-4	Sequence 4, Appl
39	63.4	3.4	14147	US-10-172-086-51	Sequence 51, Appl
40	63.2	3.4	1223197	US-10-027-632-179264	Sequence 179264,
41	63	3.4	2053	US-10-017-128-3	Sequence 3, Appl
42	63	3.4	159095	US-10-017-128-3	Sequence 3, Appl
43	62.6	3.4	1678	US-10-155-533-6	Sequence 6, Appl
44	62.6	3.4	1910	US-10-155-533-2	Sequence 2, Appl
45	62.6	3.4	5314	US-10-155-533-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-10-239-676-90/c
: Sequence 90, Application US/10239676
: Publication No. US20030082609A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIEPENBROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
: FILE REFERENCE: 5013.1003
: CURRENT APPLICATION NUMBER: US/10/239, 676
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: PCT/EP01/03968
: DE 10019058.8
: DE 10019173.8
: DE 10032529.7
: DE 10043826.1
: PRIOR FILING DATE: 2001-04-06
: 2000-04-06
: 2000-04-07
: 2000-06-30
: 2000-09-01
: NUMBER OF SEQ ID NOS: 228
: SEQ ID NO 90
: LENGTH: 5689
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
: NAME/KEY: unsure
: LOCATION: (1818, 1930..1931, 1934..1935, 1947..1948, 1966)
: US-10-239-676-90

Query Match 4.4%; Score 81.6; DB 14; Length 5689;
Best Local Similarity 45.2%; Pred. No. 0.00062;
Matches 400; Conservative 0; Mismatches 477; Indels 7; Gaps 3;
QY 907 TTTTATACAGCAAGCAAAATTTTAAACAAAGAAATTTTCAGCAGCAAAAT 966

Db	2074	TTTTCCCTCGGAAAAAAAAAAAAAAAAATATATAAAAAAAAAAACATATATAAAAAAAAAAAAAAAAAATTT	2015
Oy	967	ATTTCGCAATTTTTTAAAGCGATPCTTGTACCGCACTTTGGCATTTTAAACCTGA	10266
Db	2014	ATTTTATATATTTTTTTTTTTTTTATATTTTAAATTAATTAATAAAAAAAAAATTAATAAATTT	19555
Oy	1027	CTATCTTATATAGTATATAGATATATCCGTTAGATTTATAAGTATGTTAAACACAGTAA	10866
Db	1954	TAAATAAATTTAAAAAANNTANNATATAAAAAAAAAATATATATATATAAATAATTTT	18955
Oy	1087	AAACATATACCTATATATTTTAAATTCGTGAATTAATTTGACAGTGATTAATTAATATTA	11466
Db	1894	AATATAAAAAATATAAATATAAAAAAATATAATTAATTAATTAATTAATTAATTAATATAT	18353
Oy	1147	AGAGTATATCTATTAAGCTTAATTAATTAATCTAAAAAAGAGTAATATATGATTTGCTA	12066
Db	1834	TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	17755
Oy	1207	----TTTAAAAAGCATAGAAAAATGAATAGAACATATTAATAAAAAAGCGTGATATCA	12626
Db	1774	ATATATATAAATAAAAAATATATATAATTAATTAATTAATTAATTAATTAATTAATTA	17155
Oy	1263	ATCTTTCTTAATATCTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	13222
Db	1714	AAAAATAAAAAAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	16555
Oy	1323	TGGCGATTTTAAACCTAAATTTATCATTTGATGCAAGTTTGATCTTCTTGATGTTGA	13822
Db	1654	ATTAACAAAAAACAACAAAAACATATATCTATCAAAAAA--AATATAATTAATTA	15957
Oy	1383	GCACATTTGGCTGGGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	14422
Db	1596	AAATATATTAACAAATTTTATTAATAAAAAAATTAATTAATTAATTAATTAATTAATTA	15372
Oy	1443	AGCTAAACCTGAATTTGATTTTAAATAATTCGCTTTTATTTATCTTGATGCAATTTGGGA	15022
Db	1536	ATATATTAATTAATTAATAAACACATTCGATACACAATTTGATTAATTAATTAATTAATTA	14777
Oy	1503	AGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	15622
Db	1476	AATATTAATTAATAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	14177
Oy	1563	TTACACATATGAGAAATTAACCATATGATTAATTAATAAGAAATATATGATTTGCA	16222
Db	1416	ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	13572
Oy	1623	ATAACAAAGATATTAATAAGAAAGTCAGCTTTAAAAAATTCGATTCGATGATGTTT	16822
Db	1386	TAAATATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	12972
Oy	1663	AAAGTAGATGA--AAATTTATTTGATGAAGAGATTTGCACTAAAGATTAATTAATTAATTA	17411
Db	1296	AAAAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	12372
Oy	1742	GTATGAATAATCTTATTTGATGATCCAGATGATGCGATTAAGAAAT	1785
Db	1236	ACTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1193
RESULT 2			
US-10-239-676-210/c			
Sequence 210, Application US/10239676			
Publication No. US20030082609A1			
GENERAL INFORMATION:			
APPLICANT: OLEK, Alexander			
APPLICANT: PIESENROCK, Christian			
APPLICANT: BERLIN, Kurt			
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation			
FILE REFERENCE: 5013 1003			
CURRENT APPLICATION NUMBER: US/10/239,676			
CURRENT FILING DATE: 2002-09-24			
PRIOR APPLICATION NUMBER: PCT/EP01/03968			

Query	Match	Similarity	Score	DB	Length
DE 10019058.8	DE 10019173.8	45.6%	77.6	14	11812
DE 10032529.7	DE 10032529.7	0	0.003	2	2412
DE 10033826.1	DE 10033826.1	0	0.003	2	2412
PRIOR FILING DATE: 2001-04-06					
2000-04-06					
2000-04-07					
2000-06-30					
2000-09-01					
NUMBER OF SEQ. ID NOS: 228					
SEQ ID NO 210					
LENGTH: 11812					
TYPE: DNA					
ORGANISM: Artificial Sequence					
FEATURE:					
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)					
US-10-239-676-210					
Query Match	Best Local Similarity	45.6%	Pred. No. 0.003	Matches 311; Conservative	0; Mismatches 369; Indels 2; Gaps 1;
Db	1056	TTACATATATTAAGTATGTTAAAGCAGTAAACCAATTAATTTATATTTAAATTCGAA	1115		
Db	2351	TTAATATTAATTAACACAAAAAATTAATAAGCAGATCAACAACTTAACCTAAATGCAA	2412		
Qy	1116	TTAATATTAATTAACACAAAAAATTAATAAGCAGATCAACAACTTAACCTAAATGCAA	1175		
Db	2471	TTAATATTAATTAACACAAAAAATTAATAAGCAGATCAACAACTTAACCTAAATGCAA	2412		
Qy	1176	TAAATAAAGAGCTAATATATGATGTGCTTTTAAAGACATTTAGAAATGAAATGAA	1235		
Db	2411	TAAATAAAGAGCTAATATATGATGTGCTTTTAAAGACATTTAGAAATGAAATGAA	1235		
Qy	1236	ACATTTATTAATAAGAGCGTGATATCAATCTTCTTCAATATCTGATTAATCTTGAAT	1295		
Db	2351	CAACATTTATTAATAAGAGCGTGATATCAATCTTCTTCAATATCTGATTAATCTTGAAT	1295		
Qy	1296	AGATTAAGACATTAATATGATGTGCTTTTAAAGACATTTAGAAATGAAATGAA	1355		
Db	2291	TAAATAAAGAGCTAATATATGATGTGCTTTTAAAGACATTTAGAAATGAAATGAA	1355		
Qy	1356	TGAAGTTTTCATCTCTGATGTGCTTGAACACAACTTTCGCTGCTGCTTCAATTAATAT	1415		
Db	2231	TTCTATTTTCAAAATTTTAAAAAATTAATAAGATTTAATTTCTGCTTTTCAATTAAT	2174		
Qy	1416	ACAGAGACGAATGAATATTTATGATGATGATGATGATGATGATGATGATGATGATGAT	1475		
Db	2173	ACTATATTAATTAATTTCAATTTATGATGATGATGATGATGATGATGATGATGATGAT	2114		
Qy	1476	GTTCTTTATTCATTCGATTCATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT	1535		
Db	2113	CTTATTTAAACAAATTTAAACAAATTTAAACAAATTTAAACAAATTTAAACAAATTTAA	2054		
Qy	1536	TCAAGATGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGAT	1595		
Db	2053	ATAAAAAAACAACCTATTAATCTATTTAAACAAATTTAAACAAATTTAAACAAATTTAA	1994		
Qy	1596	AAAAATTAAGAGATTTATTTGATTTGGAATTAACAAAGATTAATAAGAAAGTACACTCTA	1655		
Db	1993	AAAAATTAAGAGATTTATTTGATTTGGAATTTAAACAAATTTAAACAAATTTAAACAAAT	1934		
Qy	1656	AAAAATTTCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT	1715		
Db	1933	CTATTAATAACAAATTTATTAATAATTTAAATTTTAAATTAATTTTAAATTTTAAATTAAC	1874		
Qy	1716	TTGCACTAAAGATTTGAAA 1737			
Db	1873	GTTCACTCATCAATTAACAA 1852			

```

: Sequence 51 Application US/10239676
: Publication No. US20030082609A1
: GENERAL INFORMATION:
: APPLICANT: OLEK, Alexander
: APPLICANT: PIEPENROCK, Christian
: APPLICANT: BERLIN, Kurt
: TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
: FILE REFERENCE: 5013.1003
: CURRENT APPLICATION NUMBER: US/10/239, 676
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: PCT/EP01/03968
: DE 10019058.8
: DE 10019173.8
: DE 10032529.7
: DE 10043826.1
: PRIOR FILING DATE: 2001-04-06
: 2000-04-06
: 2000-04-07
: 2000-06-30
: 2000-09-01
: NUMBER OF SEQ ID NOS: 228
: SEQ ID NO 51
: LENGTH: 9539
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-51

```

[illegible]

OY	1612	TGATTTTCGAATAACAAACAATCTTAAAGAAATCGCCTTTAAAAATATCTGCATTGA	1671
Dd	5639	AATTAATTTTAAAAACAAAATAATTAATATAAAAAAATTAADCTAAATAATTAAACAATCAA	5580
OY	1672	TCGATGTTTTAAAAGTAGATGCAAAAAATTATTTGATGAAGAGATTTCGACTTAAGATAT	1731
Dd	5579	AACAAATTTTAAAACTCAAAAAAATAATTAACCAAAACAAAACGTT---ATTAAAAACAAT	5524
OY	1732	TGAAAATATCTATGAAAATCTTATTGATGATCCAGATGATGCCATTAAGAATCACAAAT	1791
Dd	5523	AAAAAATTTTAAAACAAAAAAATTAATAATTTTAAAAAATAATTAACAAAAAACAAACAAT	5464
OY	1792	TCCGACGAGAAATGCGTATGCTTTCGCAATTTAAAAAATAACCAAGCAAAAAAGAGATC	1846
Dd	5463	TTAAATTAATTAATAATCAATATATATAACCAAAAAAAAAATTTTAAAAAATAATTAATAATC	5409

```

RESULT 4
US-10-027-632-128109/c
: Sequence 128109, Application US/10027632
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: TITLE OF INVENTION: Polymorphisms in the Human Genome
: FILE REFERENCE: 108827.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: CURRENT FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/218,006
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,483
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 128109
: LENGTH: 755
: TYPE: DNA
: ORGANISM: Human
US-10-027-632-128109

```

	Query Match	4.0%:	Score 74.4;	DB 14;	Length 755;
	Best Local Similarity	51.3%:	Pred. No. 0.006;		
	Matches 199;	Conservative 1;	Mismatches 182;	Indels 6;	Gaps 1;
Oy	967	ATTTTGGAAATTTTTTAAGGCGATCTGTGCACGCCACTTTTGCCATATTAAAACTGA	1026		
Db	752	AATTTTAAATATATATATATATATATAAATATATGTGATTCCTATATATATATATA	693		
Oy	1027	CTACTCTTAATAGTAAAGTATATACCGTAGATTTTAAAGTATGTTAAAACGACGAA	1086		
Db	692	TATTTAT	633		
Oy	1087	AAACAATPACTTATATATTTAATTCCTGAATTAATTTGCAGCATTTTAAATATATA	1146		
Db	632	ATATCCTATTATATGATATATATATATATATATATATATCTATTAATATATATATATA	573		
Oy	1147	AGACATATATPCTATTAGCTTAAATATAPAACTFAAAAAAGGTAAATATATGCTTGTA	1206		
Db	572	TATATATATATCTATTA	513		
Oy	1207	TTTAAAAAACGATTAGAAATGAATAGAACAT-----TATAAAAAGCGGTGATATC	1260		
Db	512	TATATATATATATATATATATATATATTTTAAATATTTAAATATATATATATATATA	453		

OY		1261	AAACCTTCTTCACAACTGCTGCATCTTGATATAGTAAGAAGATTAAAGCGATGA	13202
Dd		452	TAAATATATAAATAAT	393
OY		1321	TGTGGCAGATTTATTAACCAATTTAT	1348
Dd		392	AKATTATAGATTATATATATAAATATAT	365
RESULT 5				
US-10-239-676-160/c				
Sequence 160, Application US/10239676				
Publication No. US20030082609a1				
GENERAL INFORMATION:				
APPLICANT: OLEK, Alexander				
APPLICANT: PIEPENBROCK, Christian				
APPLICANT: BERLIN, Kurt				
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation				
FILE REFERENCE: 5013.1003				
CURRENT APPLICATION NUMBER: US/10/239.676				
PRIORITY FILING DATE: 2002-09-24				
PRIOR APPLICATION NUMBER: PCT/EP01/03968				
DE 10019058.8				
DE 10019173.8				
DE 10032529.7				
DE 10043826.1				
PRIOR FILING DATE: 2001-04-06				
2000-04-06				
2000-04-07				
2000-06-30				
2000-09-01				
NUMBER OF SEQ ID NOS: 228				
SEQ ID NO 160				
LENGTH: 9515				
TYPE: DNA				
ORGANISM: Artificial Sequence				
FEATURE:				
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)				
US-10-239-676-160				
Query Match 4.0%; Score 74.4; DB 14; Length 9515;				
Best Local Similarity 46.7%; Pred. No. 0.0093;				
Matches 279; Conservative 0; Mismatches 311; Indels 8; Gaps 1.				
OY		1137	TAAATATTAAGACATATATCTATTAGCTTAAATATTAACCTAAAAAGAGTAATATAT	11968
Dd		3001	TAAATATTAATTAATATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTA	2942
OY		1197	GGATTGCTATTTAAAAGCATTAAGAAATGAATAGAACATATATAAAAAAGAGGTGA	12565
Dd		2941	AAAAAAAAAAAAAAAAAAAAAACAAGGAAAAAACAACAAAAAATAATATCTTAT	2882
OY		1257	TATCAATCTTCTTCCATATCTTGCACTACTTGATATAGATTAAGCATTAATGCTGA	1316
Dd		2881	TACTTAATTAATTAATTAATTAATTTCTTCTTATATATATATATATATATATATATAT	2822
OY		1317	TGAATGTGGCATATTTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1376
Dd		2821	AAACCGCGCTACGCGCTATATAAATATATTTGATTCGACCTCTTTCTTCTTAA	2762
OY		1377	TGTTGAGCACAAATTTGGCTGGCTTCATTAATTAATTAATTAATTAATTAATTAATTAAT	1429
Dd		2761	CAAAGCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2702
OY		1430	AATATTTATCATCAGCTAAACGAAATGATTTTTAAANAATCTGCTTTATTCAT	1488
Dd		2701	AAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2642
OY		1489	TCGATCAATTTGGAGAATTTAATTTAGCATGAGATGATACATCAAGAAATGATT	1548
Dd		2641	ACAAATATTTCAATTAATTTCTTAATAAATAACCAAAATTAACATTTCTTTACAAAA	2582
OY		1549	AAAGATTAGTAATTAACATATGAGAAATAACCATATGAAATTAATTAATTAATTAAGAT	1608

QY	DB	Sequence	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
QY	Db	1609 ATATTGATTTTCGAAATATAACAAGATATAAAGAAGTCAGCTCTTAAAAATATCTGCAT	4.0%: Score 73.2; DB 14; Length 8085;	43.7%: Pred. No. 0.014;	0;	0;	0;	0;	0;
QY	Db	2521 AAAAAAAAAATTAATATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			0;	0;	0;	0;	0;
QY	Db	1669 TGATCGATGTTTAAAGTAGATGAAAAATTTATGTGAAGAGATTTGCACATAA			0;	0;	0;	0;	0;
QY	Db	2461 AAAAAAAAAATTAACCAAAAAAAAAAAAAAAAAAATATTACAAATTAATCTTACTATAA			0;	0;	0;	0;	0;
QY	Db	1271 TACAAATCGTCACTTACTTGTGATATAGATTAAGAATTAAGATTAAGATGATGAGCAAT			0;	0;	0;	0;	0;
QY	Db	6288 AAAAAATCTTCCACATTTAAATTCCTCTTCTCCAAAAACAATCTTCAAAAATATTTAAAT			0;	0;	0;	0;	0;
QY	Db	1331 TTATTAATCTCAAAATTTATCCATGATGAGAAATTTGATCTTCTGATGTTGAGCACAAT			0;	0;	0;	0;	0;
QY	Db	6228 TAACCCGATCGACACTTAATTTATTAATCAATAATTTTATTCATCAAAAAATTCATCAATTT			0;	0;	0;	0;	0;
QY	Db	1391 TCGGCTGGGCTTTCATTAATAATTAATACAGAGACAATGAATATTATCATCAGCTATAA			0;	0;	0;	0;	0;
QY	Db	2581 ATAAATATACATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			0;	0;	0;	0;	0;
QY	Db	1609 ATATTGATTTTCGAAATATAACAAGATATAAAGAAGTCAGCTCTTAAAAATATCTGCAT			0;	0;	0;	0;	0;
QY	Db	2521 AAAAAAAAAATTAATATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			0;	0;	0;	0;	0;
QY	Db	1669 TGATCGATGTTTAAAGTAGATGAAAAATTTATGTGAAGAGATTTGCACATAA			0;	0;	0;	0;	0;
QY	Db	2461 AAAAAAAAAATTAACCAAAAAAAAAAAAAAAAAAATATTACAAATTAATCTTACTATAA			0;	0;	0;	0;	0;
QY	Db	1271 TACAAATCGTCACTTACTTGTGATATAGATTAAGAATTAAGATTAAGATGATGAGCAAT			0;	0;	0;	0;	0;
QY	Db	6288 AAAAAATCTTCCACATTTAAATTCCTCTTCTCCAAAAACAATCTTCAAAAATATTTAAAT			0;	0;	0;	0;	0;
QY	Db	1331 TTATTAATCTCAAAATTTATCCATGATGAGAAATTTGATCTTCTGATGTTGAGCACAAT			0;	0;	0;	0;	0;
QY	Db	6228 TAACCCGATCGACACTTAATTTATTAATCAATAATTTTATTCATCAAAAAATTCATCAATTT			0;	0;	0;	0;	0;
QY	Db	1391 TCGGCTGGGCTTTCATTAATAATTAATACAGAGACAATGAATATTATCATCAGCTATAA			0;	0;	0;	0;	0;

Accession	Sequence	Position
Db	AAAAAATTCGATACCAAAATACTAAATACCCCAACATCAAAAATTATAAAAATATATA	6109
Db	6168	
Oy	1451 CTGAAATGATTTAAAAAATACGCTGTTCTTATTATTCATGCATCAATGGAGACATTTA	1510
Oy	6108 ATTTTAAACAACATCAACATCTTATATATTTTCCCATCAAAATTTTCAATCTTTT	6049
Db	1511 ATTACGATGAGATGATACANTACATCAAGAAATGATTAAGATTAATACATATATACACAT	1570
Oy	6048 AAAAACAACGCCCCCAACATCTTACACAAAAAATTTATATACCAACCAAAATCAAAAC	5988
Db	1571 ATGGAGAAATPACCATATGAATTAATAAATTAAGAATATATTTGATTTGAAATPACAA	1630
Oy	5988 AAATATATATTTATCAAAATTTAAATTAATAAACCAATACCAATTTTACACACGACGAA	5929
Db	1631 AGATATAAAGAA 1644	
Oy	5928 AATCAAAATTTATATA 5915	

```

RESULT 7
US-10-027-632-113786
: Sequence: 113786, Application US/10027632
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: FILE REFERENCE: 108827.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: CURRENT FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/218,006
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,463
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 113786
: LENGTH: 3252
: TYPE: DNA
: ORGANISM: Human
: US-10-027-632-113786

```

[illegible][illegible]

```

RESULT 8
US-10-027-632-113787
Sequence 113787, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027.632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 113787
LENGTH: 3252
TYPE: DNA
ORGANISM: Human
Query Match 3.9%; Score 72.6; DB 14; Length 3252;
Best Local Similarity 44.9%; Pred. No. 0.014;

```


[illegible][illegible]

2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 96
LENGTH: 15732
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-96

Query Match 3.9%; Score 72; DB 14; Length 15732;
Best Local Similarity 44.9%; Pred. No. 0.024;
Matches 311; Conservative 0; Mismatches 380; Indels 1; Gaps 1;

QY 1012 ATATTAAACCGACTATCTTATTAAGTATATACGATATACCGTTAGATTAAAGTAT 1071
DB 5130 ATATATTAATTAAGAAATATGATTAAGCAATTAAGTAAATTTATATTTAG 5189
QY 1072 GTTAAACGAGTAAACATTAATTAATTAATTTCTGATTAATTAATTTGACAGTGA 1131
DB 5190 AGAAATTAATTAAGAAATTAATTAATTTATTAATTAAGTAAAGATATGATGAGG 5249
QY 1132 TATTTAATTAATTAAGATATATCTATTAAGCTTAATTAATTAAGTAAAGAGTAA 1191
DB 5250 GATATTTAATTAAGATTAATTAATTAATTAAGAAAGATTTTAATTAATTAATTTGAT 5309
QY 1192 TATATGATTTGATATTAATTAATTAAGCAATTAAGAAATGAATTAAGACATTTAAGAAAGAC 1251
DB 5310 TTTATTTAATTAAGAAATTTAAAGTAAATTAATTAATTAAGTAAAGTAAATTTAAAGAA 5369
QY 1252 GGTGATATCAATCTTCTTACATCTTCTTACATCTTCTTACATCTTCTTACATCTT 1311
DB 5370 GGAATAATTAATTAAGCAATTAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAA 5428
QY 1312 GGTGATATCAATCTTCTTACATCTTCTTACATCTTCTTACATCTTCTTACATCTT 1371
DB 5429 AAAAAAGAGTGTGTTTGAATAATGATTAATTAATTAAGTAAATTTTAATTTGAATTTGA 5488
QY 1372 CTGATGTTGAGCACAATTTGGCTGGCTTCAATTAATTAAGCAATTAAGCAATTAAGCA 1431
DB 5489 TATTAAGAAAGAAAGAGCAATTTAAATTAATTAAGTAAATTAAGTAAATTAAGTAA 5548
QY 1432 TATTTAATTAATTAAGCAATTAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAA 1491
DB 5549 AATAGATTTTATTAAGAAATTAAGCAATTTTGAATTAATTAATTAAGTAAATTTGA 5608
QY 1492 ATCAATTTGGAAGCAATTTAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAA 1551
DB 5609 ATTTAATTAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTTGA 5668
QY 1552 GGTATGATTAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAA 1611
DB 5669 TTAATAAGAAATTAAGAAATTTGAAATTAAGTAAATTAAGTAAATTTGAAG 5728
QY 1612 TTGATTTGAAATTAAGAAATTAAGTAAATTAAGTAAATTAAGTAAATTTGCAATTTGA 1671
DB 5729 TTAATAATTTTATTAAGAAATTAAGTAAATTAAGTAAATTTGATTAATTTGA 5788
QY 1672 TCGATGTTTAAAGTAAATTAAGTAAATTTATTT 1703
DB 5789 AATATTAATAAGAAATTAAGTAAATTTT 5820

RESULT 14
US-10-239-676-143/C
Sequence 143, Application US/10239676
Publication No. US20030082609A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
FILE REFERENCE: 5013.1003
CURRENT APPLICATION NUMBER: US/10/239,676
CURRENT FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: PCT/EP01/03968
DE 10019058.8
DE 10019173.8
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-04-06
2000-04-06
2000-04-07
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 228
SEQ ID NO 143
LENGTH: 8801
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-143

Query Match 3.7%; Score 69; DB 14; Length 8801;
Best Local Similarity 46.7%; Pred. No. 0.065;
Matches 356; Conservative 0; Mismatches 400; Indels 7; Gaps 4;

QY 910 TTATACGAGCAAGCAAAACAAATTAATTAAGCAAGAAATTTCCAGCAAGCAAAATATT 969
DB 2305 TTTAAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2246
QY 970 TTGAAATTTTAAAGCGGATCTGCTACCGCACTTTGCAATTTTAAACCTGAGTA 1029
DB 2245 CTATATTTATTTACTATCTTCTCTATATATATATATATATATATATATATATATATAT 2186
QY 1030 TCTTATTAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTA 1089
DB 2185 TCTTAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2126
QY 1090 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1149
DB 2125 AATTAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2068
QY 1150 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1209
DB 2067 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2009
QY 1210 AAAAAAGCAATTAAGAAATTAAGCAATTAATTAAGCAATTAATTAAGCAATTAATTA 1269
DB 2008 ATACTTCATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1949
QY 1270 TTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1329
DB 1948 TACTTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1889
QY 1330 ATTATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1389
DB 1888 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1829
QY 1390 TTGCGCTGGCTTCAATTAATTAATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1449
DB 1828 ATTAACATTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1769
QY 1450 ACTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1509
DB 1768 ATACCTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1709
QY 1510 AATTAGAT-GAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1565
DB 1708 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1649
QY 1566 CACATATGCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1625

DB 1648 CACAAATACCTTAACCTTCTAAATATCTATTAATTAATTAATCTCTTAAATATA 1589
QY 1626 ACAAAGATATATAAGAAAGTCAGCTCTTAAAAATATCTGCAT 1668
DB 1588 AAAAAACATATAACGATATACCTTTTATTAATAATCATCTTTAT 1546

RESULT 15

US-10-239-676-18/C
; Sequence 18, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 18
; LENGTH: 5979
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-239-676-18

Query Match 3.7%; Score 67.8; DB 14; Length 5979;
Best Local Similarity 45.4%; Pred. No. 0.094;
Matches 358; Conservative 0; Mismatches 427; Indels 4; Gaps 3;

QY 1043 ATAGATATATCCGTAGATTAATAAGTATGTAATAAGAGTAATAACATTAATAT 1102
DB 3900 ATATATATTTTTTTTTTCCAAATCCCTTAAATTCCTTCTATTAATTAATTTCTTT 3841
QY 1103 ATTTATCTGATTAATTAATTTGACAGTGATTAATTAATTAATTAAGAGTATATCA 1162
DB 3840 AAAAAACTTAATAAAATTTCTACATTAATAATTAATAATCCAAATATTTCTATTAAT 3781
QY 1163 GCTTAATATATACTAAAAAAGGTAAATATATATATGATTTGATTTAAAAAGCATTA 1222
DB 3780 ACTTCCAAAAATTTTAATTTTAATTTTATTTTACTTTTATATACATAATATA 3721
QY 1223 AAAATGAATAGACATTAATAAAAAAGAGGTGATATCAATCTTTCTACATATCTGC 1282
DB 3720 ATATTAATATTAATAATTAATAAACTTAACCTTCACTAAACTTTTATAATAACAAT 3661
QY 1283 ATTACTTGATTAATGAATGAAGCATTAATGATGAATGTCGATTTTAAACTCAA 1342
DB 3660 ATATCTATTAATATCAAAAACTTTATTAATTTAACTTAACGATCAATCTTAACAAAAATATA 3601
QY 1343 ATTTATCATGATGAAGTTTGTATCTTCTGATTTGAGC-ACAATTTGGCGGGCT 1401
DB 3600 TATATATTTTACATACACTTTACATTTTCAATTAATAAACTATTAACAAATATAAAT 3541
QY 1402 TTCAATTAATATATACAGAGAGAAATGATATTTATCATCAGCTAATAAGTAAATGAT 1461
DB 3540 TTATATATATAAATAATATCTCAAAATATAATATTTTCTATTTAAAAAATTAATA 3481
QY 1462 TTTAAATAATCTGCTTCTTATATCATGATGATGGAAGATTTAATGATGAG 1521
DB 3480 CCACCAAGCTTTTCTTTTATTTTATTTTTCATCAA--ACACCAATCAACTTAATAAAT 3423

QY 1522 ATGACTACAAATACATCAAGAAATGATTAAGAGTTAGATTAATTAACATATGAGAAATA 1581
DB 3422 TCAATTAACCTTACCACTAAAAAATCAAAATTAACAAATCTACTATTAATAAATAA 3363
QY 1582 ACCATATGATTAATTAATTAAGAGATATATGATTTGGAATTAACAAAGTATTAAG 1641
DB 3362 AATTAATTAATTAACAAAAATCTTATTAATTAATTAACAAACAAAAATTTCTTTAACA 3303
QY 1642 AAGTCA-GCTCTTAATAATATCTGATTCATGATGATTTTAAAGTAGATGAATAATTT 1700
DB 3302 AATTCATATCAATTAATAATAAATTAATTTCTTACCAACCAATTAACATTAATTT 3243
QY 1701 ATTTGATGAAGAGATTTGCACTAAAGATTAATTAATGAAATATGCTATGAAATCTAT 1760
DB 3242 AACTTCTTAATAAATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3183
QY 1761 GATCCAGATGATGCAATGAAGAAATCACAATTCGACGAGAAATGCTATGCTTCCGC 1820
DB 3182 AACCAAAATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 3123
QY 1821 ATTAATAA 1829
DB 3122 CTCTTAATA 3114

Search completed: July 17, 2003, 07:01:01
Job time : 691.973 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 17, 2003, 02:19:35 ; Search time 89.0057 Seconds
(without alignments)
6363.996 Million cell updates/sec

Title: US-08-876-132-2
Perfect score: 1847
Sequence: 1 AGATCTCAACAGCTTTAA...AAAAACAGCAAGAGATCT 1847

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6C_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/6D_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74.6	4.0	19124	2	US-08-487-826B-13
2	65.8	3.6	20674	4	US-09-641-638-651
3	65.2	3.5	3095	6	5231168-1
4	64	3.5	1431	4	US-09-316-083-2
5	63.2	3.4	7218	1	US-08-332-463-14
6	62.4	3.4	19124	2	US-08-487-826B-13
7	62.2	3.4	6243	2	US-09-056-075-1
8	61.6	3.3	665	2	US-08-883-795A-36
9	61.4	3.3	615	4	US-08-998-416-186
10	60.2	3.3	5852	1	US-07-867-106-2
11	59.4	3.2	636	4	US-08-998-416-1137
12	59	3.2	6124	4	US-08-213-419B-3
13	58	3.1	837	4	US-08-998-416-288
14	58	3.1	8920	2	US-08-446-855A-1
15	58	3.1	920	4	US-09-150-741-1
16	57.8	3.1	4140	3	US-08-894-731-2
17	57.6	3.1	872	4	US-08-998-416-487
18	57.2	3.1	6124	4	US-08-213-419B-3
19	56.4	3.1	658	4	US-08-998-416-595
20	55.6	3.0	677	4	US-08-998-416-1092
21	55.6	3.0	821	4	US-08-998-416-541
22	54.8	3.0	2739	4	US-09-134-001C-156
23	54.4	2.9	5852	2	US-08-883-795A-36
24	54.4	2.9	5852	1	US-07-867-106-2
25	54.4	2.9	20674	4	US-09-641-638-651
26	54	2.9	636	4	US-08-998-416-1137
27	54	2.9	8133	1	US-08-480-604A-5

28	54	2.9	8133	2	US-08-405-496A-5	Sequence 5, Appl1
29	54	2.9	8133	4	US-08-915-136-5	Sequence 5, Appl1
30	54	2.9	8133	4	US-08-957-310-5	Sequence 5, Appl1
31	53.6	2.9	1689	1	US-07-991-867B-41	Sequence 41, Appl1
32	53.6	2.9	1689	2	US-08-544-332-41	Sequence 41, Appl1
33	53.6	2.9	1689	4	US-09-370-861A-41	Sequence 41, Appl1
34	53.6	2.9	1947	4	US-09-370-861A-74	Sequence 74, Appl1
35	53.6	2.9	8457	1	US-07-991-867B-1	Sequence 1, Appl1
36	53.6	2.9	8457	2	US-08-544-332-1	Sequence 1, Appl1
37	53.6	2.9	8457	4	US-09-370-861A-1	Sequence 1, Appl1
38	52.8	2.9	290	3	US-08-617-860B-13	Sequence 13, Appl1
39	52.8	2.9	1511	1	US-07-991-867B-8	Sequence 8, Appl1
40	52.8	2.9	1511	1	US-08-107-755A-8	Sequence 8, Appl1
41	52.8	2.9	1511	2	US-08-544-332-8	Sequence 8, Appl1
42	52.8	2.9	1511	4	US-09-370-861A-8	Sequence 8, Appl1
43	52	2.8	711	4	US-09-134-001C-1758	Sequence 1758, Ap
44	52	2.8	2435	4	US-09-306-593-1	Sequence 1, Appl1
45	51.6	2.8	1956	4	US-08-559-896B-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knodde Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487, 826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13
Query Match 4.0%; Score 74.6; DB 2; Length 19124;
Best Local Similarity 46.8%; Pred. No. 1.9e-06;

[illegible][illegible]

DB 980 CAGATGGTCTTTTATCATCTATGTAATCCCAAGATACATTATTTAAANA 1039
1785 TCACATTCG 1794
DB 1040 TGAGACCTAG 1049

RESULT 5

US-08-232-463-14/C

Sequence 14, Application US/08232463

Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHEFFELINGER, F.

APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: PTZgpt-F15

US-08-232-463-14

Query Match 3.4%; Score 63.2; DB 1; length 7218;

Best Local Similarity 16.0%; Pred. No. 0.00031;

Matches 91; Conservative 224; Mismatches 252; Indels 1; Gaps 1;

QY 1277 ACTGCATTACTTGTATAGTAATTAAGCATTAATGATGATCAATGTGGCGATTATATA 1336
DB 1643 ACCTGCAGGAATCTTATATAGCATGAAAAACAAATGAAATTCATATATATTTTA 1584
QY 1337 ACTCAATTTATCCATGATGAAAGTTTGAATCTTGTAGTGTAGACAAATTCGGCT 1396
DB 1583 CATACATATATCTTAATATGAAGTGTGATGTGACTAGCGTACGCTCTAGAC 1524
QY 1397 GGGCTTTCAATTAATATACAGACAAATGAATATTTATCATCAGCTAAACTGAA 1456
DB 1523 GCATCTATTCAGTTCAAAAAACGCAATGTAGCATCACTGAATTAACCTATCTATGCA 1464

QY 1457 ATGATTTTAAAAATATCGTCTTATATTCATTCGATTCGAGAAATTAATACG 1516
DB 1463 GT-AGTTAAAGATTAAGAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRR 1405
QY 1517 ATGAGATGCATGATACATCAAGAAATGATTAAGATTAATTAATTAACATATGAG 1576
DB 1404 RRR 1345
QY 1577 AATTAACCATATGATTAATTAAGATATATATGATTTTCGAAATTAACAAAGAT 1636
DB 1344 RRR 1285
QY 1637 AAAAGAACTGCTCTTAAATATCTCATTCATGATGATTTTAAAGATGAGAAA 1696
DB 1284 RRR 1225
QY 1697 ATTTATGATGAAGATTTTCAACTAAGATATGAAATATGATGAAATTCAT 1756
DB 1224 RRR 1165
QY 1757 TGATGATCAGATGATGCATTAAGAAATACAAATTCGACAGAGAAATGCTATGCTT 1816
DB 1164 RRR 1105
QY 1817 CCGATTAATAAACAGCAAGCAAAAGAGA 1844
DB 1104 RRR 1077

RESULT 6

US-08-487-826B-13/C

Sequence 13, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitlus, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobe Martens Olson & Beat

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelien, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 19124 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

;; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
;; NUMBER OF SEQUENCES: 39
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BERESKIN & PARR
;; STREET: 40 King Street West
;; CITY: Toronto
;; STATE: Ontario
;; COUNTRY: Canada
;; ZIP: M5H 3Y2
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/883,795A
;; FILING DATE: 27-JUN-1997
;; CLASSIFICATION: 435
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gravelle, Micheline
;; REGISTRATION NUMBER: 40,261
;; REFERENCE/DOCKET NUMBER: 7841-062
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (416) 364-7311
;; TELEFAX: (416) 361-1398
;;
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 665 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; IMMEDIATE SOURCE:
;; CLONE: Rh 32
;; US-08-883-795A-36

Query Match 3.38; Score 61.6; DB 2; Length 665;
Best Local Similarity 46.7%; Pred. No. 0.00043;
Matches 196; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

QY 926 ACAAAATAAATAAAGAAAGAAATTTGACGAGAAATAATTTTGAATTTTAAAG 985
DB 447 ATAAATATGTAATTAATAACATTTAATTATAAATATGTAATTAATAACATTTAATT 388
QY 986 GCGATACCTTGCTACCGCATTTTGCATATTTAAACCTGACTATCTTTAAGTTAATA 1045
DB 387 ATAAATATGTAATTAATAACATTTAATTATAAATATGTAATTAATAACATTTAATT 328
QY 1046 GATATATCGGTAGATTATAAGTANGTAAAGCGATAAAGCAATTAACCTATATATT 1105
DB 327 ATAAATATGTAATTAATAACATTTAATTATAAATATGTAATTAATAACATTTAATT 268
QY 1106 TAAATTCGATTAATTTGACGAGTAATTTAATAATATAAGATATATTCATTTGCT 1165
DB 267 ATAAATATTTAATTAATAATTTAATTATAAATATGTAATTAATAACATTTAATT 208
QY 1166 TAAATATACATAAAGAGAGTAATATATGATGCTGATTTAAAGCAATTAAGAA 1225
DB 207 ATAAATATTTAATTAATAATTTAATTATAAATATGTAATTAATAACATTTAATT 148
QY 1226 ATGAATATAGACATTTATAAAGAGCGGATATCAATCTTCTTCAATCACTTGCAAT 1285
DB 147 ATAAATATTTAATTAATAATTTAATTATAAATATGTAATTAATAACATTTAATT 88
QY 1286 ACTTGATATAGATTAAGCGATTAAGGCGATGAGATGCGATATATTAACCAAT 1345
DB 87 ATAAATATTTAATTAATAATTTAATTATAAATATGTAATTAATAACATTTAATT 28

RESULT 9
US-08-998-416-186/c

;; Sequence 186, Application US/08998416
;; Patent No. 6239264
;;
;; GENERAL INFORMATION:
;; APPLICANT: Philippsen, Peter
;; APPLICANT: Pohlmann, Rainer
;; APPLICANT: Steiner, Sabine
;; APPLICANT: Mohr, Christine
;; APPLICANT: Wendland, Jurgen
;; APPLICANT: Knechtle, Philipp
;; APPLICANT: Reibischung, Corinne
;;
;; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPYII
;; TITLE OF INVENTION: AND USES THEREOF
;; NUMBER OF SEQUENCES: 1152
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: No. 6239264artis Corporation
;; STREET: 3054 Cornwalis Road
;; CITY: Research Triangle Park
;; STATE: No. 6239264th Carolina
;; COUNTRY: USA
;; ZIP: 27709
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/998,416
;; FILING DATE: 24-DEC-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: CH 0016/97
;; FILING DATE: 31-DEC-1996
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Weigs, J. Timothy
;; REGISTRATION NUMBER: 38,241
;; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8587
;; TELEFAX: 919-541-8689
;;
;; INFORMATION FOR SEQ ID NO: 186:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 615 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: PAG1074RP
;; US-08-998-416-186

Query Match 3.38; Score 61.4; DB 4; Length 615;
Best Local Similarity 46.8%; Pred. No. 0.00047;
Matches 269; Conservative 0; Mismatches 296; Indels 10; Gaps 2;

QY 1061 TTATTAAGTATGTAATAAAGAGATTAACAATACTATATATTATTTCTGCAATATA 1120
DB 608 TTATTAATATTTAATTAATTAATTAATAAGATTAATAATTAATTAATTAATTAATTA 549
QY 1121 TTGACAGCATTAATTAATATATTAAGATATATCTATTAGCTTAATTAATTAATTA 1180
DB 548 TT--AAATTAATTAATTAATAAATGAATATATTAATTAATTAATTAATTAATTAATTA 491
QY 1181 AAGAGGTAATATATGATGTTGCTATTTAAAGCAATTAAGAAATGAATATGACATTT 1240
DB 490 TAAGAAATTTAAGTTAAATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTA 431
QY 1241 ATAAAGAGCGGATATCAATCTTCTTCAATCACTTGCAATTTGATTTAGATATA 1300
DB 430 AATTAACATTAATTTTAAAGATATTAATTAATTAATTAATTAATTAATTAATTAATTA 371
QY 1301 AAGCATTAATATGATGATATGTCGATATTAATTAATTAATTAATTAATTAATTAATTA 1360
DB 370 AATTAATTAATCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 311

QY	1013	TATTTAAACCGACTACCTTTATAGTAAATAGTAAATACCGGTAGTAAATGAAGTATG	1072
Db	636	TTTTATAGTATTTTTTACCTACATCTTTATATATATTTTATTAATTAATTAATATGAT	577
QY	1073	TTAAAAACGACTAAAAACAATTAATTAATTTAAATTCGAATTAATATTTGACAGTAT	1132
Db	576	AAATATTTAATTAATATTAATTTAATTTAATTAATTAATTAATTAATTAATTAATTAAT	517
QY	1133	TATTTAATTAATTTAAGACATTAATCTATTACTTAAATATTAACATAAAAAAGAGTAAT	1192
Db	516	TATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAATTAATTA	457
QY	1193	ATATGATCTTGCTATTTTAAAAAACACATTAAGAAATGAACATTAATTAATAAAGACG	1252
Db	456	TTCTTATTAATAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	400
QY	1253	GTGATATCAATCTTTCTTACATTAATCTGATCTGATTAATGAATTAAGACATTAATG	1312
Db	399	ATAATTAATAAATAATTAATTTACAAATTTTAAAAATTTAAATCTTTAATTAATTAATAA	340
QY	1313	GTGATGAATGGGGATTAATTTAATTAACCAATTTATTCATGATGAAGAAAGTTTGAAT	1370
Db	339	ATTATTTTAAATAACATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	280
QY	1371	TCCTGATGTGGAGCACAAATTTGGCTGGGCTTTCAATTAATAATTAATTAACAGACAGCAATGA	1430
Db	279	TATTAAGAAATATATATATATCTAATTAATTTAATTAATTAATTAATTAATTAATTAATTAAT	220
QY	1431	ATATTTATCACTACACTAAACTGAAGATTTTAAAAAATATACGCTTTTATTAATCAATC	1490
Db	219	GACATTAATTAATGAT - ATTCATATTTAATTAATTAATTTAATTAATTAATTAATTAATTAATTAAT	162

QY 1457 ATGATTTTAAAAA

[illegible]

DB 1970 TTAAATATATATATATATATAGAGATTTTCTCTCATTTTGTGAACATTTTAAGATTT 1911
QY 1517 ATGAGATGATACATACATACAGAAATGATTAAGATTTGATATATACATATGAG 1576
DB 1910 CGGATATCAAT 1851
QY 1577 AAATACCATATGATTAATATATATATATATATATATATATATATATATATAT 1636
DB 1850 TTGAT 1791
QY 1637 AAAGAAAGTACGCTTAAATATATATATATATATATATATATATATATATATAT 1696
DB 1790 TAAAGATATACAAAGGATATATATATATATATATATATATATATATATATAT 1733
QY 1697 ATTTATGATGAGAGATTTGCAACTAAGATATATGATTAAGAAATCTAT 1756
DB 1732 ATTAATATTTAAATTTGATCTCATATATATATATATATATATATATATATAT 1673
QY 1757 TGATGATCCAGATGATGCAATAGAAAT 1785
DB 1672 AAAAAAAAAAAATTTGTAATATATAT 1644

RESULT 13

US-08-998-416-288/c

Sequence 288, Application US/08998416

Patent No. 6239264

GENERAL INFORMATION:

APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reibschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/998,416

FILING DATE: 24-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: CH 0016/97

FILING DATE: 31-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 288:

SEQUENCE CHARACTERISTICS:

LENGTH: 837 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: PAG1241RP
US-08-998-416-288

Query Match 3.1%; Score 58; DB 4; Length 837;

Best Local Similarity 45.3%; Pred. No. 0.0024;

Matches 251; Conservative 0; Mismatches 300; Indels 3; Gaps 1;

QY 946 AAAAAATTTTCGACGCAAAAAATATTTTGAATTTTAAAGCGATCTGCTACCGCACT 1005
DB 631 AAGATTTTAACTACATCTTAAATATATATATATATATATATATATATATATAT 572
QY 1006 TTGCGATATTTAAACCTGACTATCTTATATAGTATATATATATATATATATAT 1065
DB 571 TTAATTTAAT 512
QY 1066 AAGTATTTAAACGATTAACAAATATATATATATATATATATATATATATATAT 1125
DB 511 AAT 452
QY 1126 CAGTGATTTAT 1185
DB 451 ATAAAGATTTAAT 392
QY 1186 GGTAT 1245
DB 391 AAT 332
QY 1246 AAAGCGGTAT 1305
DB 331 AATAAT 275
QY 1306 TTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1365
DB 274 AAGAAAT 215
QY 1366 GATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1425
DB 214 AATAGTAT 155
QY 1426 AATGAT 1485
DB 154 AGTAAT 95
QY 1486 CATTGATCAATG 1499.
DB 94 AATGAT 81

RESULT 14

US-08-446-855A-1/c

Sequence 1, Application US/08446855A

Patent No. 5849573

GENERAL INFORMATION:

APPLICANT: Stewart, Thomas S

APPLICANT: Flores, Maria V

APPLICANT: O'Sullivan, William J

TITLE OF INVENTION: Nucleotide sequence encoding carboxymyl

TITLE OF INVENTION: Phosphate synthetase II

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon & Vanderyhe PC

STREET: 1100 No. 5849573th Glebe Road, 8th Floor

CITY: Arlington

STATE: Virginia

COUNTRY: USA

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.24

CURRENT APPLICATION DATA: US/08/446, 855A


```

1      FILING DATE: 06-Jul-1995
2
3      CLASSIFICATION: 435
4
5      ATTORNEY/AGENT INFORMATION:
6
7      NAME: Mitchell, Leonard C
8      REGISTRATION NUMBER: 29,005
9
10     REFERENCE/DOCKET NUMBER: 4
11
12     TELECOMMUNICATION INFORMATION
13
14     TELEPHONE: 703-816-4000
15
16     TELEFAX: 703-816-4100
17
18     INFORMATION FOR SEQ ID NO: 1:
19
20     SEQUENCE CHARACTERISTICS:
21
22     LENGTH: 8920 base pairs
23     TYPE: nucleic acid
24     STRANDEDNESS: single
25     TOPOLOGY: linear
26
27     MOLECULE TYPE: genomic
28
29     OS-08-446-855A-1

```

Query Match	3.1%;	Score 58;	DB 2;	Length 8920;
Best Local Similarity	46.8%;	Pred. No. 0.0035;		
Matches 249; Conservative	0;	Mismatches 280;	Indels 3;	Gaps 2

[illegible]

RESULT 15
US-09-150-741-1/c
Sequence 1, Application US/09150741
Patent No. 6183996
GENERAL INFORMATION:
APPLICANT: Stewart et al.
TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
Patent No. 6183996
TITLE OF INVENTION: Synthetase II
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/150,741
CURRENT FILING DATE: 1998-09-10
EARLIER APPLICATION NUMBER: PL6380
EARLIER FILING DATE: 1992-12-16

```

: EARLIER APPLICATION NUMBER: A093/000617
: EARLIER FILING DATE: 1993-12-02
: EARLIER APPLICATION NUMBER: 08/446,855
: EARLIER FILING DATE: 1995-07-06
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 8920
: TYPE: DNA
: ORGANISM: Plasmodium falciparum
US-09-150-741-1

```

Query Match	3.1%	Score 58	DB 4	Length 8920
Best Local Similarity	46.8%	Pred. No.	0.0035	
Matches 249	Conservative	0	Mismatches 280	Indels 3
				Gaps 2

QY	923	AAAACAAAATAATAACAAAGAAAATATTTTCGAGCCGAAAAATATTTGGAAATTTT	982
Dd	563	AAATTTAAATAAATATATATATTAATATTTATAGAACACAAATATTTGGAGATATTA	504
QY	983	AAGCGATCTCTGCTACCGCCTTTTGGCAATTTTAAAAACCTGACATCTTTATAGTTA	1042
Dd	503	TATTCGTATATATATATATAGATGGTTTATATTAAGAGCTAAACCTATTTATGTGT	444
QY	1043	ATAGCTATATCCGTAGATTTATAAGTATGTTTAAAGACGAGTAAACCAATATACCTATAT	1102
Dd	443	ACACATCAAAATATTTAAATATTTGATTCATATTAATTTGAAATATATTAATATAGTTTAT	384
QY	1103	ATTATATTCGAATATATATTTTGACAGTGATTTTATATATATTAAGACATATATCTATTA	1162
Dd	383	ATTTCAGATATTTTATATTA-ACAAATGAAACACAAACATATATATATATATATATAT	325
QY	1163	GCTTAATATATACATAAAAAAGAGTAAATATATGATTTGCTATTTTAAAAAGCATATG	1222
Dd	324	ATATATATATATATATATATATTAATAATCTTAATATGATTTGTTAAATATATTAAGAAAA	265
QY	1223	AAAAATGAATATGAACATTATAAAAAGAGGATATATCAAAATCTTTTACAAATACCTTGC	1282
Dd	264	AAAAAAGAAAAAGTTTTTTATCTATGTTATATATATGAATATATCATATATATATACATC	205
QY	1283	ATTACTTTGATATATGAATTAAGCATTAATAATGATGAAATGCGGATATTTTAAACTTAA	1342
Dd	204	ATATCAAAATATAAATATTTTTTATCTTTTATTTTTTATTTATTTATTTTTCCTTCAA	145
QY	1343	AT--TTATCCATGATGAAGATTGATCTTCTGATGGTTGAGCAAAATTTGCGGCGGC	1400
Dd	144	ATGCTATATAAATTTTACTCTTCTCTCGTTAAGCCTTCTTAAGCCCTTCACAGTATCG	85
QY	1401	TTTCAATTAATATATACGAGACGCAAAATGAATATTTATCATCAGCTAAACT	1452
Dd	84	AAATTTCTACCTCANTCAAGACGAGTGCATATGTTTCTTTCTTTAAAT	33

Search completed: July 17, 2003, 05:12:46
Job time : 94.0057 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2003, 19:52:25 ; Search time 4292.71 Seconds

(without alignments)
11254.112 Million cell updates/sec

Title: US-08-876-132-1

Perfect score: 1660

Sequence: 1 AGATCTACACAGCAAAATT.....AAGGACTCCGACGACATCT.1660

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Genembl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_dr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1308.6	78.8	3661	1	AF128889	AF128889 Pantoea C
2	109	6.6	6282	6	AX251493	AX251493 Sequence
3	106.2	6.4	108908	3	PFMAL3P8	AL034560 Plasmid
4	104.2	6.3	349980	6	AX344555	AX344555 Sequence
5	101.6	6.1	131682	9	AL672277	AL672277 Human DNA
6	101.6	6.1	231912	2	AC087566	AC087566 Mus muscu
7	99.6	6.0	174384	2	AC009524	AC009524 Homo sapi
8	98.4	5.9	130540	2	AC079417	AC079417 Mus muscu
9	96.6	5.8	14867	3	AE001398	AE001398 Plasmid
10	96.6	5.8	84563	5	AC096885	AC096885 Danio rer
11	93.6	5.6	6644	6	E23356	E23356 Virus vecto
12	93.6	5.6	7372	6	E23357	E23357 Virus vecto
13	93.6	5.6	7797	6	E23355	E23355 Virus vecto
14	93.6	5.6	7996	6	E23359	E23359 Virus vecto
15	93.4	5.6	268147	2	AC116966	AC116966 Dictyoste
16	91.8	5.5	67970	3	PFMAL1P3	AL031746 Plasmid
17	91.4	5.5	191366	2	AC115351	AC115351 Rattus no
18	90.8	5.5	257703	3	CEY111B2A	AL132904 Caenorhab
19	90.2	5.4	14006	6	AX346860	AX346860 Sequence
20	90	5.4	11000	2	CEY111B2_0	Z98857 Caenorhabdl
21	89.8	5.4	300695	2	AC079431	AC079431 Mus muscu
22	89.8	5.4	6668	6	AX346599	AX346599 Sequence
23	89.8	5.4	16544	2	AC125605	AC125605 Homo sapi
24	88.4	5.3	83234	2	AC034104	AC034104 Homo sapi
25	87.6	5.3	5420	3	AF056936	AF056936 Plasmid
26	87.6	5.3	7442	6	AX251441	AX251441 Sequence
27	87	5.2	170568	2	AC107605	AC107605 Rattus no
28	86.6	5.2	15548	6	AX347057	AX347057 Sequence
29	86.4	5.2	303091	2	AC084799	AC084799 Mus muscu
30	86.2	5.2	156060	2	AC004153	AC004153 Plasmid
31	86.2	5.2	169794	2	AC004688	AC004688 Plasmid
32	86	5.2	224448	2	PFMAL4P4	AL035477 Plasmid
33	85.4	5.1	9716	2	AC115603	AC115603 Dictyoste
34	85	5.1	66441	3	PFMAL1P4	AL031747 Plasmid
35	85	5.1	215287	2	AC079432	AC079432 Mus muscu
36	84.6	5.1	147962	2	AC130632	AC130632 Rattus no
37	84.2	5.1	82469	2	AC116330	AC116330 Dictyoste
38	84	5.1	143585	2	AC013349	AC013349 Homo sapi
39	83.8	5.0	9539	6	AX277889	AX277889 Sequence
40	83.8	5.0	9539	6	AX323566	AX323566 Sequence
41	83.6	5.0	7571	6	AX345429	AX345429 Sequence
42	83.6	5.0	7823	6	AX278035	AX278035 Sequence
43	83.6	5.0	7823	6	AX323804	AX323804 Sequence
44	83.6	5.0	7823	6	AX344489	AX344489 Sequence
45	83.6	5.0	7823	6	AX346963	AX346963 Sequence

ALIGNMENTS

RESULT 1
AF128889
LOCUS
DEFINITION Pantoea citrea strain ATCC31623 plasmid pZG500, complete plasmid
ACCESSION AF128889
VERSION AF128889.1 GI:4457229
KEYWORDS
SOURCE Pantoea citrea.
ORGANISM Pantoea citrea.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Pantoea.
REFERENCE 1 (bases 1 to 3661)
AUTHORS Billie M. and Delic V.
TITLE Isolation and characterization of a cryptic plasmid from *Erwinia*

Pred. No. is the number of results predicted by chance to have a

FEATURES	source
JOURNAL	citreus ATCC 31623
MEDLINE	J. Appl. Microbiol. 83 (4), 485-492 (1997)
PUBMED	98012514
REFERENCE	9351229
ATTORNS	2 (bases 1 to 3661)
TITLE	Balic Nezic,M. and Delic,V. Sequence Analysis and Functional Characterization of Replication Region in Plasmid pz6500 from Pantoea citrea ATCC 31623
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 3661)
AUTHORS	Balic Nezic,M. and Delic,V. Direct Submission
TITLE	Submitted (16-FEB-1999) Biotechnology, PRIVA d.d., Research Institute, Piliž baruna Filipovica 25, Zagreb 10 000, Croatia Location/Qualifiers
FEATURES	1. .3661
source	/organism="Pantoea citrea" /strain="ATCC31623" /db_xref="ATCC:31623" /db_xref="taxon:53336" /plasmid="pz6500" complement(join(3188..3661,1..177)) /note="RNAII" 99..317 /note="Orf1" /codon_start=1 /transl_table=1 /product="unknown" /protein_id="AAD21205.1" /db_xref="GI:4457230" /translation="MGOAQLCWSAHAPAYSIGALAKLKHKNLIGLGFLYOKKKIK 177 OKIKFERKNILFEFGDGYRTFAIFKT" /note="OrfV" complement(334..342) /note="putative" /bound_moiety="DnaA" 458..472 /bound_moiety="HF" 489..704 /note="Orf2" /codon_start=1 /transl_table=1 /product="unknown" /protein_id="AAD21206.1" /db_xref="GI:4457231" /translation="MDCVFKALENEIEHYKKDGIKSPLOYLHYFDIDKALNGDECG DIINSNLSIDESLIFLMLSQASGISK" complement(845..1066) /note="Orf3" /codon_start=1 /transl_table=1 /product="unknown" /protein_id="AAD21207.1" /db_xref="GI:4457232" /translation="MPSSGSSIGFSDIFNIFSCSSSSINFSSTLKHRSQIFLRAD FLYLILFRNYQILLEYSGYGSICVIL" complement(1193..1348) /note="Orf4" /codon_start=1 /transl_table=1 /product="unknown" /protein_id="AAD21208.1" /db_xref="GI:4457233" /translation="MODLFCWHLFFLHPITYTHHFLFLYIKKFSMAKFLISTEAENI SHIFLKK" 1991..2365 /note="Orf5" /codon_start=1 /transl_table=1 /product="putative heat-shock protein" /protein_id="AAD21209.1" /db_xref="GI:4457234"
CDS	
misc_RNA	
CDS	
rep_origin	
protein_bind	
protein_bind	
CDS	
CDS	
CDS	
CDS	

CDS	<p>/translation="MDFRSRLTJLNEKKDLDEKIAESELKARKJGTGPGVLEWMTME MKRINLDVNEETAGQYRKLEFNKNVSHKSDSLVTGLLEGCTRNSPDKTRSAFRCIC ERLQWMDIEFTTKIILQYKEKO" 2362..2991 /function="putative resolution of catenane plasmid dimers prior to plasmid segregation to the daughter cells" /note="Orf6" /codon_start=1 /transl_table=11 /product="lambda-integrase like protein" /protein_id="AAD21210.1" /db_xref="GI:4457235" /translation="MKRADTMDIDJFRRLKNKSTYDRYAGFLSICSTTCGRPAEVLKG IEIERNRDEGISIFKILGAKVGNDRGSEPTLHPDLHYHNEQMYLISOLKONKEF KYPKCKLYNSROYLYIOHRTFSLYTLRRHRAVSLKSGADDFITAFPLGRVYQSD FYGVARSKSGIANTGVCGSDYVANKNSQFAVSSTPQISLKSHPKK"</p>									
	-35_signal	1	3154..3159							
	-10_signal	1	3175..3180							
	misc_RNA		3188..3288							
			/note="RNA1"							
	-10_signal	1	3295..3300							
	-35_signal	1	3311..3316							
BASE COUNT	1366 a	472 c	730 g	1093 t						
ORIGIN										
Query Match	78.8%	Score 1308.6;	DB 1;	Length 3661;						
Best Local Similarity	91.1%	Pos. No. 5e-182;								
Matches 1510;	Conservative 0;	Mismatches 12;	Indels 136;	Gaps 6;						
QY	8	CACAAAGCAAAATTGAAAAAATAGATAAAA--TTTTCGACAGTATTAAAGCCGACTTAAAC	66							
Db	1433	CACAAAGCAAAATTGAAAAAATAGATAAAAATTTTTCGACAGTATTAAAGCCGACTTAAAC	1492							
QY	67	AAATGAGTGAAGAAGAAAGAAAAAATAATATACA--TATTTGAGTTAGTAAAGAGAA	124							
Db	1493	AAATGAGGAGAGAGAGAAAGAAAAATTAAATACACTATTGAGAGTTAGTAAAGAGAA	1552							
QY	125	AGAAAAAATAAGAAAGACTCGGCTTACAGTCGAAAAACGAGAAATTAATAAAAAGAA	184							
Db	1553	AGAAAAAATAAGAAAGAA--CTCGGCTTACAGTCGAAAAACGAGAAATTAATAAAAAGAA	1611							
QY	185	GAGACTGAGATTTTATATGAAAAATCGAGAGAAAAAATTTAAATTTTCATTTTCGAG	244							
Db	1612	GAGACTGAGATTTTATATGAAAAATCGAGAGAAAAAATTTAAATTTTCATTTTCGAG	1671							
QY	245	GGATTTAAATTTGTGTAGATTGATGAAAAATCTAGATAAAAAATGACATCAAAAATGTG	304							
Db	1672	GGATTTAAATTTGTGTAGATTGATGAAAAATCTAGATAAAAAATGACATCAAAAATGTG	1731							
QY	305	TGATTTGACATTTTGAATTAACGTACTATATATCAATTAATGGGGTTGTCTATTTATTT	364							
Db	1732	TGATTTGACATTTTGAATTAACGTACTATATATCAATTAATGGGGTTGTCTATTTATTT	1791							
QY	365	TGCGAAGATTTGAAAATCTGAGTGAAGAAAAATACTTTGCGAGAGCAAAAAAACCCCTTGC	424							
Db	1792	TGCGAAGATTTGAAAATCTGAGTGAAGAAAAATACTTTGCGAGAGCAAAAAAACCCCTTGC	1851							
QY	425	GTTTTTTTCAATGACCTTTGGAAAAAATTAATCTATGTGACGGGTACGAAACCTTGAATTT	1911							
Db	1852	GTTTTTTTCAATGACCTTTGGAAAAAATTAATCTATGTGACGGGTACGAAACCTTGAATTT	1911							
QY	485	TTTTCATTGGAAGATTTGAAAAAATAAGGCAAAAAAGAACTCAAAATGAAAAAATTTATTA	544							
Db	1912	TTTTCATTGGAAGATTTGAAAAAATAAGGCAAAAAAGAACTCAAAATGAAAAAATTTATTA	1971							
QY	545	TAAAAAAGAGATCGGATATGATTTTAAAGCAGAAAACTGCATTTGAATGAAAAAA	604							
Db	1972	TAAAAAAGAGATCGGATATGATTTTAAAGCAGAAAACTGCATTTGAATGAAAAAA	2031							
QY	605	AGATTGGAAAAAATCTATGCTGAGAGTGAATTTAAAGCAAAAAAATTGGAACTCAAC	664							
Db	2032	AGATTGGAAAAAATCTATGCTGAGAGTGAATTTAAAGCAAAAAAATTGGAACTCAAC	2091							

LOCUS	DEFINITION	AX251493	Sequence 461 from Patent WO0168912.	6292 bp	DNA	linear	PAT 05-0CT-2001
Oy		665	CGGGTGTCTTTTACGAATGACGATGGAAGAATGAAATGAAATATATCAACCTCGATGTTAA	724			
Db		2092	CGGTGTGTCTTTTACGAATGACGATGGAAGAATGAAATGAAATATATCAACCTCGATGTTAA	2151			
Oy		725	TGAAGAAACACACAGGCTCAATATAGGAATATTCTCAAAATATTAAGTTGAGCTAGTAATC	784			
Db		2152	TGAAGAAACACACAGGCTCAATATAGGAATATTCTCAAAATATTAAGTTGAGCTAGTAATC	2211			
Oy		785	AGATGATCTAGTAACGGGACTATTTAGAGCTGTGAACTCGAAATAGTTTGTATTAACAAG	844			
Db		2212	AGATGATCTAGTAACGGGACTATTTAGAGCTGTGAACTCGAAATAGTTTGTATTAACAAG	2271			
Oy		845	AAGGCCCTTCGTTTTTGTATTGTGTGAGAGAAATTCAGCACTGAGAAAAAGAGCTGATTA	904			
Db		2272	AAGGCCCTTCGTTTTTGTATTGTGTGAGAGAAATTCAGCACTGAGAAAAAGAGCTGATTA	2299			
Oy		905	TGCAAGAAAGTAAGAAAGATTCGATACAAATGAAGCAAAAACCTAAGAGGCTTTGAAAT	964			
Db		2300	TGCAAGAAAGTAAGAAAGATTCGATACAAATGAAGCAAAAACCTAAGAGGCTTTGAAAT	2299			
Oy		965	GAGTTTGTCTTTTGATTAAGATTTTTTGAGTGAAATATAGAAATTCAGATGATGATTT	1023			
Db		2300	GAGTTTGTCTTTTGATTAAGATTTTTTGAGTGAAATATAGAAATTCAGATGATGATTT	2322			
Oy		1024	-CTCAACAACAAAAAGACTCTGCAAGTAATAAAGAAAAACATGAAGAAAGGAGACACAATG	1082			
Db		2323	-CTCAACAACAAAAAGACTCTGCAAGTAATAAAGAAAAACATGAAGAAAGGAGACACAATG	2382			
Oy		1083	GATGATATATTTTAAAGAGGCTTAAAAAATATATTAATCTACATATGATTCGTTATGCTGATTC	1142			
Db		2383	GATGATATATTTTAAAGAGGCTTAAAAAATATATTAATCTACATATGATTCGTTATGCTGATTC	2442			
Oy		1143	CTTTCTATTTGTCGATTACAGGTTGCAGACACAGACAGAAGTTTAAAGGTAATAGATA	1202			
Db		2443	CTTTCTATTTGTCGATTACAGGTTGCAGACACAGACAGAAGTTTAAAGGTAATAGATA	2502			
Oy		1203	GTAAGAAACAGATATGAGGATGATATCTTTTAAATCTTTGGTGCAAGGTTGGAAT	1262			
Db		2503	GTAAGAAACAGATATGAGGATGATATCTTTTAAATCTTTGGTGCAAGGTTGGAAT	2562			
Oy		1263	GACAGAGGCGCAAGAGGAAGAAACATTAATTTGATTATCAAAATATCATGATATAG	1322			
Db		2563	GACAGAGGCGCAAGAGGAAGAAACATTAATTTGATTATCAAAATATCATGATATAG	2622			
Oy		1323	CAATATGATATATATTTTGTCCGCAATTAAGAATATATAATTTTCTCAACACAGATGG	1382			
Db		2623	CAATATGATATATATTTTGTCCGCAATTAAGAATATATAATTTTCTCAACACAGATGG	2682			
Oy		1383	AAGCTCTACACAGAGCTTGAGGCAATPACTCTACATCCACATPAGAACGTTTCACTGAT	1442			
Db		2683	AAGCTCTACACACAGCTTGAGGCAATPACTCTACATCCACATPAGAACGTTTCACTGAT	2742			
Oy		1443	ACACTTGTGTCACAGGTTGGAGAGATGCTCAAGGATCCCGGTGCAAGATTCACCATTA	1502			
Db		2743	ACACTTGTGTCACAGGTTGGAGAGATGCTCAAGGATCCCGGTGCAAGATTCACCATTA	2802			
Oy		1503	GCGGCTTTTTTGGGTCCACAGAGTACTCAAGACGAGAGGATTAACGGCTATGCTGCTTC	1562			
Db		2803	GCGGCTTTTTTGGGTCCACAGAGTACTCAAGACGAGAGGATTAACGGCTATGCTGCTTC	2863			
Oy		1563	GTCGNAAGAGTGATTCGCTGTAACTGTTGAGTCTCTGATGTTGTGAAGCAAAACA	1622			
Db		2862	GTCGNAAGAGTGATTCGCTGTAACTGTTGAGTCTCTGATGTTGTGAAGCAAAACA	2922			
Oy		1623	GAGTCAGTTTNGCTGTATCCAAAGGACTCCGAGCCCAATCT	1660			
Db		2922	GAGTCAGTTTNGCTGTATCCAAAGGACTCCGAGCCCAATCT	2959			

[illegible]

Query Match	Best Local Similarity	Score	DB	Length
Matches 307; Conservative	50.7%;	0; Mismatches 293; Indels 5; Gaps 2;		108908;
gene				
CDS				
CECELVHNNDNDPEMRKVMQFHDRTTORROEDELQERQVCKTCNEKEIKIL				
KDKLEKEINEKFAFLQTDIHSDAIPICITCKTSMADKVEKCTKOTQWUGIIVASSGV				
LAGIEGALVAMKPNALQOTAEALKAABDIIIVGEGAGKVKVIGGLDALELEKLI				
GSMPTPELEGICIVKSLASIIYEKROTGICASTJSDKTCEDIGISITMWDGTY				
GAPPTPELVINGIVGKTEADVADEARESESTNKNVOETRLBAGNSSISSIN				
ASIAIVVILIIMVITFIYILRYPRKKIKKKIKYIKLLEF				
join(61248..61301,61548..62459)				
/gene="PFC0040w/MAL3P8.7"				
join(61248..61301,61548..62459)				
/gene="PFC0040w/MAL3P8.7"				
aa; Similarity to other P.falciptarum rifins. Member of a				
multi-gene family, predicted signal and transmembrane				
sequence; predicted using hexon				
/codon_start=1				
Query Match	6.4%;	Score 106.2;	DB 3;	Length 108908;
Best Local Similarity	50.7%;	Pred. No. 6.1e-07;		
Matches 307; Conservative	0; Mismatches 293; Indels 5; Gaps 2;			
47	TATTAAAGCCACATTTAAACAAATGAGAGAGAAAGAAAAATTAATACATTTT	106		
Db	39088 TATTCGAAAAAAGAAAAATGAAAAAAGAAAAAAGAAAAAAGAAATATTAACAAT	3914		
QY	107 TGAGTTAGTAAAGAGAAAGAAAAATTAAGAAGACCTCGCTTAACAGTCGAAAAAC	166		
Db	39148 TTATTTAAATATTAATAAAAAAAGAAAAAAGAAAAATATTTTAAATAATTAATAAT	3920		
QY	167 AGAATATATTAAGAAAGAGACGCTGATTTTATATGAAATCGTGAGGAAAAAATTT	226		
Db	39208 TATATATATAAATTAATAATTTTGAATTAATAAATAAATAATTTATCAAAAAAATTT	3926		
QY	227 TTAATTTTCATTTTCGAGGATTAATTTGTTGTAAGTTGATGAAAAATCTGATATAAAA	286		
Db	39268 AAAAAAATTTTATATATTAATAAAAAATGATTTATTAATAAAAAATTTATAGAAATAAAA	3932		
QY	287 TGCAGATCAAAA-ATGTTGGAATTTGACATTTTGAATPACGTATATCAATAATGG	345		
Db	39328 AAAACAAAAAAGAAAAATTAATGTTAAAAAATAATATATCATTAATAATAAAAA	3938		
QY	346 GGGTTTGCTTTTATTTTTCGCAAGATTTGAAATCTGAGTGAAGAAATAGTTGCGA	405		
Db	39388 ATGGAAGAAATTTTAAATATGAAAAATTAATTAATAATAATAATAATAATTTATA	3944		
QY	406 GAGCAAAAAAACCCTGCGCTTTTTCCAATGACTTGAAAAATTCATCTGAGCGG	465		
Db	39448 AAAATTAATAA-----TGAAAAATTTTATCAAAAAAATTTAAAAAATTTATATTAATA	3950		
QY	466 TAGCGAAACTTTGAAATTTTTCACATTTGAAATTTGAAAAATTAAGCAAAAGAACTCA	525		
Db	39504 AATTAATGATTTTAATAAAAAATTTATTTGAAATTAATAATAATAATAATAATAATCTTA	3956		
QY	526 AATGAAAAAATATATTTATTAATAAAAAAGAGATCGGATATGATTTTAAAAAGCAAAAC	585		
Db	39564 AAAAAAATTTATATTAATAATAAAAAATTAATAATTTAAAAAATTTAAAAAATTTAAAA	3962		
QY	586 TGACATTCGAATGAAGAAAGATTTGAAAAAATCTATGCTGAGAGTGAATTAAGCAAA	645		
Db	39624 AAAAAATTAATAAAAAAATTTAAAAAATAATAATAATAATAATAATAATAATAATAA	3968		
QY	646 AAAAA 650			
Db	39684 ATAAA 39688			
RESULT 4				
AX344555/c	349980 bp	DNA	linear	PAT 01-FEB-2002
LOCUS	Sequence 6 from Patent WO0200932.			
ACCESSION	AX344555			
VERSION	AX344555.1	GI:18492441		
FEATURES				
synthetic construct.				

COMMENT On Aug 22, 2002 this sequence version replaced gi:21911559.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL, SWISSPROT, TrEMBL, WP, WormPeP, information on the WormPeP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX> RP11-309M23 is from the library RPCR-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES
Source Location/Qualifiers

1. 131682
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-309M23"
/clone_1lb="RPCR-11.2"
BASE COUNT 34805 a 31378 c 30235 g 35264 t
ORIGIN

Query Match 6.1%; Score 101.6; DB 9; Length 131682;
Best Local Similarity 46.6%; Pred. No 2.7e-06;
Matches 360; Conservative 0; Mismatches 409; Indels 3; Gaps 1;

16 AATGTAATAATGATTAATTTTCGACGATTTAAAGCCGACTTAAACCAATGAGTG 75
118367 AAAAAATTAATGAAATATGAAATTAATTAATTAATTAATTAATTAATTAATTA 118426
76 AACAAGAAAAAGAAAAATTAATTAATTTGAGTTAGTAAAGAAAGAAAAATTA 135
118427 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118486
136 AAGACAGCTCGGCTTAACAGTGAAGAACCAAGAAATTAATTAATTAATTAATTAATTA 195
118487 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118546
196 TTTTAATGCAATCGTGAAGAAAAAATTTTAAATTTTCATTTTCGAGGATTAATTTG 255
118547 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118603
256 TTGTAGTGTGATCAAAATCTAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 315
118604 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118663
316 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 375
118664 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118723
376 AATATCTGAGTGAAGAAATTAATTTGCGAGCAAAAAAACCCCTTGGCTTTTTCATTA 435

Db 118724 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118783
Qy 436 ATGACTTTGCAAAAAATTCATTGTGACGGTAGCAAACTTTGAAATTTTTCATTGCA 495
Db 118784 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118843
Qy 496 AATTTGAAAAATTAAGCAAGAACTCAAAATGAAAAATTAATTAATTAATTAATTAATTA 555
Db 118844 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118903
Qy 556 GATCGATATGATTTTAAAGCAAGAACTGACATTAATTAATTAATTAATTAATTAATTAATTA 615
Db 118904 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 118963
Qy 616 AATCTATGCTGAGAGTGAATTAAGCAAAAAATTTGGCACTCAACCCGTTGTTT 675
Db 118964 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 119023
Qy 676 TAGAAATGACGATGAAGAAATGATGAAGAAATTAATTAATTAATTAATTAATTAATTA 735
Db 119024 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 119083
Qy 736 CAGCTCAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 787
Db 119084 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 119135

RESULT 6
AC087566 231912 bp DNA linear HTG 10-JAN-2001
LOCUS AC087566
DEFINITION Mus musculus clone Rp23-127H3, WORKING DRAFT SEQUENCE, 57 unordered pieces.
AC087566
VERSION AC087566.1 GI:12061432
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 231912)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Mouse
JOURNAL Unpublished
2 (bases 1 to 231912)
REFERENCE DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT

-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 0
Center clone name: RPCR-23_127H3

Summary Statistics
Consensus quality: 150164 bases at least Q40
Consensus quality: 175115 bases at least Q30
Consensus quality: 186307 bases at least Q20
Estimated insert size: 216300; agarose-fp estimation
Estimated insert size: 226312; sum-of-contigs estimation
Quality coverage: 3.5 in Q20 bases; agarose-fp estimation
Quality coverage: 3.34 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently consists of 57 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 1074: contig of 1074 bp in length
* 1075 1174: gap of unknown length
* 1175 2270: contig of 1096 bp in length
* 2271 2370: gap of unknown length
* 2371 3511: contig of 1141 bp in length
* 3512 3612: gap of unknown length
* 3612 4783: contig of 1172 bp in length
* 4784 4884: gap of unknown length
* 4884 6377: contig of 1494 bp in length
* 6378 6478: gap of unknown length
* 6478 7712: contig of 1234 bp in length
* 7712 7812: gap of unknown length
* 7812 9092: contig of 1281 bp in length
* 9093 9193: gap of unknown length
* 9193 10330: contig of 1138 bp in length
* 10331 10430: gap of unknown length
* 10431 11677: contig of 1247 bp in length
* 11678 11777: gap of unknown length
* 11778 13186: contig of 1409 bp in length
* 13187 13287: gap of unknown length
* 13287 14565: contig of 1279 bp in length
* 14566 15787: gap of unknown length
* 15788 15887: contig of 1122 bp in length
* 15888 17012: contig of 1125 bp in length
* 17013 17112: gap of unknown length
* 17113 18353: contig of 1241 bp in length
* 18354 18453: gap of unknown length
* 18454 19611: contig of 1158 bp in length
* 19612 19711: gap of unknown length
* 19712 21084: contig of 1373 bp in length
* 21085 21184: gap of unknown length
* 21185 22398: contig of 1214 bp in length
* 22399 22498: gap of unknown length
* 22499 23629: contig of 1131 bp in length
* 23630 23729: gap of unknown length
* 23730 25112: contig of 1383 bp in length
* 25113 25212: gap of unknown length
* 25213 26529: contig of 1317 bp in length
* 26530 26629: gap of unknown length
* 26630 27773: contig of 1144 bp in length
* 27774 27873: gap of unknown length
* 27874 29211: contig of 1338 bp in length
* 29212 29311: gap of unknown length
* 29313 30438: contig of 1127 bp in length
* 30439 30538: gap of unknown length
* 30539 31930: contig of 1392 bp in length
* 31931 32030: gap of unknown length
* 32031 33112: contig of 1082 bp in length
* 33113 33212: gap of unknown length
* 33213 34380: contig of 1168 bp in length
* 34381 34480: gap of unknown length
* 34481 35515: contig of 1035 bp in length
* 35516 35615: gap of unknown length
* 35616 37145: contig of 1530 bp in length
* 37146 37245: gap of unknown length
* 37246 38655: contig of 1410 bp in length
* 38656 38755: gap of unknown length
* 38756 40418: contig of 1663 bp in length
* 40419 40518: gap of unknown length
* 40519 42089: contig of 1571 bp in length
* 42090 42189: gap of unknown length
* 42190 43485: contig of 1296 bp in length
* 43486 43585: gap of unknown length
* 43586 45557: contig of 1972 bp in length
* 45558 45657: gap of unknown length
* 45658 47561: contig of 1904 bp in length
* 47562 47661: gap of unknown length
* 47662 50270: contig of 2609 bp in length
* 50271 50370: gap of unknown length
* 50371 52463: contig of 2093 bp in length
* 52464 52563: gap of unknown length
* 52564 54241: contig of 1678 bp in length

```

```

* 54242 54341: gap of unknown length
* 54342 56282: contig of 1941 bp in length
* 56283 56382: gap of unknown length
* 56383 57724: contig of 1342 bp in length
* 57725 57825: gap of unknown length
* 57826 60404: contig of 2580 bp in length
* 60405 60504: gap of unknown length
* 60505 63040: contig of 2536 bp in length
* 63041 63141: gap of unknown length
* 63142 64314: contig of 1174 bp in length
* 64315 64414: gap of unknown length
* 64415 68063: contig of 3649 bp in length
* 68064 68163: gap of unknown length
* 68164 72715: contig of 4552 bp in length
* 72716 72815: gap of unknown length
* 72816 79858: contig of 7043 bp in length
* 79859 79959: gap of unknown length
* 79960 87305: contig of 7347 bp in length
* 87306 87405: gap of unknown length
* 87406 94304: contig of 6899 bp in length
* 94305 94404: gap of unknown length
* 94405 99973: contig of 5569 bp in length
* 99974 100073: gap of unknown length
* 100074 110741: contig of 10668 bp in length
* 110742 110841: gap of unknown length
* 110842 116556: contig of 5715 bp in length
* 116557 116556: gap of unknown length
* 116557 126137: contig of 9481 bp in length
* 126138 126237: gap of unknown length
* 126238 139661: contig of 13424 bp in length
* 139662 139761: gap of unknown length
* 139762 147923: contig of 8162 bp in length
* 147924 148023: gap of unknown length
* 148024 160030: contig of 12007 bp in length
* 160031 160130: gap of unknown length
* 160131 175441: contig of 15311 bp in length
* 175442 175541: gap of unknown length
* 175542 191133: contig of 15592 bp in length
* 191134 191233: gap of unknown length
* 191234 231912: contig of 40679 bp in length.

```

FEATURES
source 1. 231912 /organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-127H3"
/lib="RPC1 mouse BAC library 23"

BASE COUNT 61588 a 52985 c 59076 g 49983 t 8280 others
ORIGIN

Query Match 6.1%; Score 101.6; DB 2; Length 231912;
Best Local Similarity 45.8%; Pred. No. 2.4e-06;
Matches 433; Conservative 0; Mismatches 497; Indels 15; Gaps 2;

```

QY 21 GAAAAAATGATTAATTTTCGCAGCTATTAAGCCGACTTAACAANTGATGAAGA 80
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8018 GAAAAAATGATTAATTTTCGCAGCTATTAAGCCGACTTAACAANTGATGAAGA 8077
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 81 GAAAAAATGATTAATTTTCGCAGCTATTAAGCCGACTTAACAANTGATGAAGA 140
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8078 GAAAAAATGATTAATTTTCGCAGCTATTAAGCCGACTTAACAANTGATGAAGA 8137
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 141 GACCTCGCTTAACGATGAGAAAAACCAGAAATATTAAGAGAGAGAGAGAGAGAG 200
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8138 AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8197
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 201 ATGGAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 260
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8198 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8257
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 261 AGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 320
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 8258 AGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8317
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

OY 321 GAAATACGATATATCAATTAATGCGGCTTGTCTATTTTATTTTCGAAGATTGAAT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8318 AGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA- 8376
OY 381 CTGAGTGAAGAAATATAGTTTGGAGACAAAAAACCTTGCCCTTTTTCCTAAATGAC 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8377 -AGAGGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 8435
OY 441 TTGGAAAAATTCATTTGAGCGGTAGCGAAACTTTGAATTTTTCATCTGGAATTT 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8436 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 8495
OY 501 GAAAAATATAGCGAAAGAACTCAATGAAAAATATTTATTTATTTAAAAAGAGATCG 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8496 AGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 8555
OY 561 GATATGATTTTAAAGCAAAACTGACATTTGATGAAAAAGAAAAAGATTTGGAAAAATC 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8556 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 8615
OY 621 TATGCTGAGTGAATTAAGCAAAAAATTTGGAACCAACCGGTGTTTATAGAA 680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8616 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 8675
OY 681 ATGACGATGAAGAAATGATGAATAATATCACTGATGTTATGAGAAACAGCAGGT 740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8676 AAAAAAGTGAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAG 8735
OY 741 CATATAGCAAAATTTATCAAAAAATTAAGTTGACATAGTAATCAGATGATCTAGTAACG 800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8736 AAAAAAGAAAAAGTATGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 8795
OY 801 GGACT-----ATTAGAGTGTGAACCTGGAACCTGATTTGATTAACCAAGAG 847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8796 GAAGGAGAGCAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAGGCTAAAAAGAAAGAG 8855
OY 848 TGCCTTTCTTTTGTATTTGTGAGAGAAATTCAGCACTGAGAAAAAGAGCTGATATGC 907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8856 AGAGGGGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 8915
OY 908 AAGAGAGTAAAAAGATTTGATTCATGAAAGCAAAACTAAGA 952
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8916 AGAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAA 8960

RESULT 7
AC009524 174384 bp DNA linear HTG 07-JUN-2000
LOCUS Homo sapiens chromosome 15 clone RP11-31701 map 15q24, LOW-PASS
SEQUENCE SAMPLING.
AC009524
AC009524.2 GI:8312269
KEYWORDS HTG: HTGS_PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 174384)
Rowen, L., Madan, A., Qin, S., Abbasi, N., Baradaran, L., Birditt, B.,
Bloom, S., Dors, M., Dickhoff, R., Fleetwood, P., Harrison, G., Kaur, A.,
Madan, A., Nesbitt, R., Shafer, T., and Hood, L.
JOURNAL Submitted (26-AUG-1999) Multimegabase Sequencing Center, University
of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT On Jun 7, 2000 this sequence version replaced gi:5776572.
----- Genome Center
Center: Multimegabase Sequencing Center

```

```

Center code: UWMSC
Web site: http://chroma.mbt.washington.edu/msg_wmw
Contact: leeuwenesystemsbiology.org
----- Summary Statistics
Sequencing vector: pUC18; 108752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
-----
NOTE: This record contains 192 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
795 794: contig of 794 bp in length
895 894: gap of unknown length
1703 1702: contig of 808 bp in length
1803 1802: gap of unknown length
2637 2637: contig of 835 bp in length
2638 2638: gap of unknown length
3557 3557: contig of 820 bp in length
3558 3557: gap of unknown length
3658 4453: contig of 796 bp in length
4454 4553: gap of unknown length
4554 5365: contig of 812 bp in length
5365 5466: gap of unknown length
5466 6259: contig of 794 bp in length
6260 6359: gap of unknown length
6360 7168: contig of 809 bp in length
7169 7268: gap of unknown length
7269 8063: contig of 795 bp in length
8064 8163: gap of unknown length
8164 8977: contig of 814 bp in length
8978 9077: gap of unknown length
9078 9872: contig of 795 bp in length
9873 9972: gap of unknown length
9973 10785: contig of 812 bp in length
10785 10884: gap of unknown length
10884 11779: contig of 795 bp in length
11779 11680: gap of unknown length
11680 12592: contig of 813 bp in length
12593 12692: gap of unknown length
12693 13490: contig of 798 bp in length
13491 13590: gap of unknown length
13591 14404: contig of 814 bp in length
14405 14504: gap of unknown length
14505 15295: contig of 791 bp in length
15295 15396: gap of unknown length
15396 16208: contig of 813 bp in length
16208 16308: gap of unknown length
16309 17112: contig of 804 bp in length
17113 17113: gap of unknown length
17113 18028: contig of 816 bp in length
18029 18128: gap of unknown length
18129 18926: contig of 798 bp in length
18927 19026: gap of unknown length
19027 19839: contig of 813 bp in length
19840 19939: gap of unknown length
19939 20738: contig of 799 bp in length
20739 20838: gap of unknown length
20838 21653: contig of 815 bp in length
21654 21753: gap of unknown length
21754 22554: contig of 801 bp in length
22555 22654: gap of unknown length
22655 23468: contig of 814 bp in length
23469 23568: gap of unknown length
23568 24368: contig of 800 bp in length

```

24369 24468: gap of unknown length
24469 25282: contig of 814 bp in length
25283 25382: gap of unknown length
25383 26185: contig of 803 bp in length
26186 26285: gap of unknown length
26286 27100: contig of 815 bp in length
27101 27200: gap of unknown length
27201 28005: contig of 805 bp in length
28006 28105: gap of unknown length
28106 28921: contig of 816 bp in length
28922 29021: gap of unknown length
29022 29827: contig of 806 bp in length
29827 29927: gap of unknown length
29927 30736: contig of 809 bp in length
30737 30836: gap of unknown length
30837 31629: contig of 793 bp in length
31630 31729: gap of unknown length
31730 32554: contig of 825 bp in length
32555 32654: gap of unknown length
32655 33452: contig of 798 bp in length
33453 33552: gap of unknown length
33553 34362: contig of 810 bp in length
34363 34462: gap of unknown length
34463 35261: contig of 799 bp in length
35262 35361: gap of unknown length
35362 36171: contig of 810 bp in length
36172 36271: gap of unknown length
36272 37067: contig of 796 bp in length
37068 37167: gap of unknown length
37168 37980: contig of 813 bp in length
37981 38080: gap of unknown length
38081 38674: contig of 794 bp in length
38675 38974: gap of unknown length
38975 39788: contig of 814 bp in length
39789 39888: gap of unknown length
39889 40686: contig of 798 bp in length
40687 40786: gap of unknown length
40787 41597: contig of 811 bp in length
41598 41697: gap of unknown length
41698 42496: contig of 799 bp in length
42497 42596: gap of unknown length
42597 43407: contig of 811 bp in length
43408 43507: gap of unknown length
43508 44317: contig of 810 bp in length
44318 44417: gap of unknown length
44419 45233: contig of 816 bp in length
45234 45333: gap of unknown length
45334 46127: contig of 794 bp in length
46128 46227: gap of unknown length
46228 47041: contig of 814 bp in length
47042 47141: gap of unknown length
47143 47937: contig of 796 bp in length
47938 48037: gap of unknown length
48038 48849: contig of 812 bp in length
48849 48949: gap of unknown length
48950 49747: contig of 798 bp in length
49748 49847: gap of unknown length
49848 50660: contig of 813 bp in length
50661 50760: gap of unknown length
50761 51559: contig of 799 bp in length
51560 51659: gap of unknown length
51660 52473: contig of 814 bp in length
52474 52573: gap of unknown length
52575 53375: contig of 802 bp in length
53376 53475: gap of unknown length
53476 54284: contig of 809 bp in length
54285 54384: gap of unknown length
54385 55187: contig of 803 bp in length
55188 55287: gap of unknown length
55288 56096: contig of 809 bp in length
56097 56196: gap of unknown length
56197 56992: contig of 796 bp in length
56993 57092: gap of unknown length

57093 57907: contig of 815 bp in length
57908 58007: gap of unknown length
58008 58805: contig of 798 bp in length
58806 58905: gap of unknown length
58906 59718: contig of 813 bp in length
59719 59818: gap of unknown length
59819 59819: gap of unknown length
60615 60714: gap of unknown length
60715 61524: contig of 810 bp in length
61525 61624: gap of unknown length
61625 62422: contig of 798 bp in length
62423 62522: gap of unknown length
62523 63332: contig of 810 bp in length
63333 63432: gap of unknown length
63433 64228: contig of 796 bp in length
64229 64328: gap of unknown length
64329 65138: contig of 810 bp in length
65139 65238: gap of unknown length
65239 66037: contig of 799 bp in length
66038 66137: gap of unknown length
66138 66948: contig of 811 bp in length
66949 67048: gap of unknown length
67049 67844: contig of 796 bp in length
67845 67944: gap of unknown length
67945 68757: contig of 813 bp in length
68758 68857: gap of unknown length
68858 69654: contig of 797 bp in length
69655 69754: gap of unknown length

Query Match 6.0% Score 99.6: DB 2: Length 174384;
Best Local Similarity 39.4%: Pred: No. 5e-06;
Matches 535: Conservative 0: Mismatches 811: Indels 13: Gaps 3;

OY 7 ACACAGCAATTCGATAAATATGATTAATTTTCGACGATTAATTAAGCCGCTTAAC 66
DB 103894 AGAAAAAATCTCTAATAAATTAAGCAGCTGTAAGAAAAAAGAAATATTA 103953
OY 67 AAATGAGTGAAG 126
DB 103954 AAAGGAAG 104013
OY 127 AAAAAATTAAG 186
DB 104014 AAAAAAATTAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 104073
OY 187 GACTGTGATTTTAATGAATTCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
DB 104074 AAACGAG 104133
OY 247 ATTAATTTGTTGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306
DB 104134 AATAAT-----AAAAAGATTAANNNNNNNNNNNNNNNNNNNNNNNNNNN 104187
OY 307 AATTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 366
DB 104188 NNN 104247
OY 367 CGAAGATTAAG 426
DB 104248 NNNNCGTTTTCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 104307
OY 427 TTTTTCATTAATGACTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
DB 104308 AAGGAGCTTCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 104367
OY 487 TACATTTGAATTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
DB 104368 AAAAAAGATTCCTTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 104427
OY 547 AAAAAAGAGATTCGATTTGAATTTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
DB 104428 AAAAAAG 104487

[illegible]

COMMENT

```

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----
Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 0
Center Clone name: RPCI-23_102A24
-----
Summary Statistics
Consensus quality: 71349 bases at least Q40
Consensus quality: 83788 bases at least Q30
Consensus quality: 89369 bases at least Q20
Estimated insert size: 93300; agarose-1p estimation
Estimated insert size: 127240; sum-of-ctrls estimation
Quality coverage: 3.69 in Q20 bases; agarose-1p estimation
Quality coverage: 2.7 in Q20 bases; sum-of-ctrls estimation..
NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1043: contig of 1043 bp in length
1
1044 1143: gap of unknown length
1144 2163: contig of 1017 bp in length
2161 2260: gap of unknown length
2261 3696: contig of 1436 bp in length
3697 3797: gap of unknown length
3797 4826: contig of 1030 bp in length
4827 4927: gap of unknown length
4927 5959: contig of 1033 bp in length
5960 6059: gap of unknown length
6060 7579: contig of 1520 bp in length
7580 7679: gap of unknown length
7680 9178: contig of 1499 bp in length
9179 9278: gap of unknown length
9279 10501: contig of 1223 bp in length
10502 10601: gap of unknown length
10602 11610: contig of 1009 bp in length
11611 11710: gap of unknown length
11711 12840: contig of 1130 bp in length
12841 12940: gap of unknown length
12941 14262: contig of 1322 bp in length
14263 14362: gap of unknown length
14363 15984: contig of 1622 bp in length
15985 16084: gap of unknown length
16085 17502: contig of 1418 bp in length
17503 17602: gap of unknown length
17603 18902: contig of 1300 bp in length
18903 19002: gap of unknown length
19003 20008: contig of 1006 bp in length
20009 20108: gap of unknown length
20109 21372: contig of 1264 bp in length
21373 21472: gap of unknown length
21473 22675: contig of 1203 bp in length
22676 22775: gap of unknown length
22776 24088: contig of 1313 bp in length
24089 24188: gap of unknown length
24189 25354: contig of 1166 bp in length
25355 25454: gap of unknown length
25455 26641: contig of 1187 bp in length
26642 26741: gap of unknown length
26742 29340: contig of 2599 bp in length
29341 29441: gap of unknown length
29441 31609: contig of 2169 bp in length
31610 31709: gap of unknown length
31710 33695: contig of 1986 bp in length
33696 33795: gap of unknown length
33796 35117: contig of 1322 bp in length

```

* 35118 35217: gap of unknown length
* 35218 38766: contig of 3549 bp in length
* 38767 38866: gap of unknown length
* 38867 43795: contig of 4929 bp in length
* 43796 43895: gap of unknown length
* 43896 45218: contig of 1323 bp in length
* 45219 45318: gap of unknown length
* 45319 50886: contig of 5568 bp in length
* 50887 57066: gap of unknown length
* 57067 57166: gap of unknown length
* 57167 64436: contig of 7270 bp in length
* 64437 64536: gap of unknown length
* 64537 74858: contig of 10322 bp in length
* 74859 86302: contig of 11344 bp in length
* 86303 86402: gap of unknown length
* 86403 96560: contig of 10158 bp in length
* 96561 130540: gap of unknown length
* 96661 130540: contig of 33880 bp in length.

FEATURES
source
1. 130540
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
/clone="RP23-102A24"
/clone_11b="RPCI mouse BAC library 23"

BASE COUNT 37619 a 23855 c 30292 g 28866 t 9888 others

ORIGIN

Query Match 5.9%; Score 98.4; DB 2; Length 130540;
Best Local Similarity 43.6%; Pred. No. 8e-06;
Matches 480; Conservative 0; Mismatches 616; Indels 4; Gaps 2;

OY 21 GAAAAATAGATTAATTTTTCGACGATTAATTAACCGCATTAACCAATGATGACAGAA 80
DB 17783 GAAAAATTAAGGAGGCGNNNGTATTAAGGCGNNNANGCGNNNANGCGNNNANGCG 17842
OY 81 GAAAAAGAAAAATTAATCATATTTTGAAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAG 140
DB 17843 AAAAAANNN 17902
OY 141 GACCTCGGCTTAACGTCGAAAAACCAGAAATATTAAGGAGGAGGAGGAGGAGGAGGAGG 200
DB 17903 AAAGAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 17962
OY 201 ATGGAATCGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 260
DB 17963 AAAAAATGTAAATTAAGATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18022
OY 261 AGTGTATGAAAAATCTAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 320
DB 18023 AAAAAAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18082
OY 321 GAAATAGCTAGATATCAATAATGCGGCTTCTATTTTATTTTTCGAGGAGGAGGAGGAG 380
DB 18083 AAAAGGAAAAATTTAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18142
OY 381 CTGAGTGAAGAAATAGTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 440
DB 18143 AGAAAAAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18202
OY 441 TTTGGAAGAAAAATCTAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 500
DB 18203 GGAAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18262
OY 501 GAAAAATTAAGGCAAGAACTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 560
DB 18263 AAGAAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18322
OY 561 GATATGATTTTAAAGCAAGAAATCTGACATTTGAT---GAAAAAGGAGGAGGAGGAGG 617
DB 18323 GAAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18382

OY 618 ATCTATGCTGAGAGTAAATTAAGCAAAAAATTTGGAGACTCAACCGGTTGTTTAA 677
DB 18383 TGAAGAAATGTAAATTAAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18442
OY 678 GAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 737
DB 18443 TAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18502
OY 728 GGTCAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 797
DB 18503 AAAAAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18562
OY 798 ACGGACATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 857
DB 18563 ATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18622
OY 858 TTTTGTATTTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 917
DB 18623 AAAAGTAAAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 18682
OY 918 AAGATTTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
DB 18683 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18742
OY 978 GATAAGATTTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1036
DB 18743 AAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18802
OY 1037 AGACTCTCAAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1096
DB 18803 AAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18862
OY 1097 GAGGCTAAAGTAAATTAAT 1116
DB 18863 TAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 18882

RESULT 9
AE001398/14867 bp DNA linear INV 16-APR-2002
LOCUS
DEFINITION
Plasmodium falciparum chromosome 2, section 35 of 73 of the complete sequence.
ACCESSION
AE001398 AE001362
VERSION
AE001398.1 GI:3845197
KEYWORDS
SOURCE
ORGANISM
Plasmodium falciparum 3D7.
Eukaryota; Alveolata; Apicomplexa; Hemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 14867)
Gardner,M.J., Tetteh,I.H., Carucci,D.J., Cummings,L.M., Arayind,L., Koonin,E.V., Shalton,S., Mason,T., Yu,K., Fujii,C., Pederson,J., Shen,K., Jing,J., Aston,C., Lai,Z., Schwartz,D.C., Pertea,M., Salzberg,S., Zhou,L., Sutton,G.G., Clayton,R., White,O., Smith,H.O., Fraser,C.M., Adams,M.D., Venter,J.C. and Hoffman,S.L. Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum
Science 282 (5391), 1126-1132 (1998)
JOURNAL
MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 14867)
Gardner,M.J.
Direct Submission
Submitted (02-NOV-1998) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20814, USA
AUTHORS
JOURNAL
TITLE
REFERENCE
2 (bases 1 to 14867)
Gardner,M.J.
Direct Submission
Submitted (02-NOV-1998) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20814, USA
FEATURES
source
1. 14867
/organism="Plasmodium falciparum 3D7"
/strain="3D7"
/db_xref="taxon:36329"
/chromosome="2"
complement(1570..2424)
/gene="PF80490C"

gene


```

ACCESSION      AC096885
KEYWORDS        AC096885.2 GI:18182756
SOURCE          HTG.
ORGANISM        Danio rerio.
                Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
REFERENCE       1 (bases 1 to 84563)
AUTHORS         Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
                Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
                Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
                Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Jin,S.-O.,
                Leaspl,R., Maduro,O.L., Maduro,V.B., Mastello,C., Maskeri,B.,
                Mastrian,S.D., McCloskey,J.C., McDowell,U., Pearson,R., Prasad,A.,
                Stantrop,S., Thomas,J.W., Thomas,P.J., Touchman,D.W.,
                Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
                Young,A., Zhang,L.-H. and Green,E.D.
                NISC Comparative Sequencing Initiative
TITLE           Unpublished
JOURNAL         2 (bases 1 to 84563)
AUTHORS         Green,E.D.
TITLE           Direct Submission
JOURNAL         Submitted (29-SEP-2001) NIH Intramural Sequencing Center, 8717
                Government Circle, Gaithersburg, MD 20877, USA
REFERENCE       3 (bases 1 to 84563)
AUTHORS         Green,E.D.
TITLE           Direct Submission
JOURNAL         Submitted (17-JAN-2002) NIH Intramural Sequencing Center, 8717
                Government Circle, Gaithersburg, MD 20877, USA
                On Jan 17, 2002, this sequence version replaced gi:15808575.
COMMENT         ----- Genome Center
                Center: NIH Intramural Sequencing Center
                Center code: NISC
                Web site: http://www.nisc.nih.gov
                Contact: nisc.mouse@nhgri.nih.gov
                ----- Project Information
                Center project name: agp
                Center clone name: J36L04

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.

FEATURES             source            Location/Qualifiers
                    1..84563
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="J36L4"
                     /clone_1lb="Incyte Genomics"
                     /misc_feature 48561..48672
                     /note="single clone coverage"
BASE COUNT        26959 a 15018 c 14630 g 27956 t
ORIGIN
Query Match      5.8% Score 96.6; DB 5; Length 84563;
Best Local Similarity 45.5%; Pred. No. 1.6e-05;
Matches 467; Conservative 0; Mismatches 549; Indels 11; Gaps 3;

32 TAAATTTTGGCAGATTATAAGCCGCATTAAACAACAAGTAGTGAGAAGAAAAGAAAA 91
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 47634 TATTAATTCCTGCTTTCTTTAAATTAATAAATCAAAATTAATGAAGAAAATTAACATAAAG 47699
92 AATAAATACATATTATTTGAGTTAGTAAAGAGAAAAGAAAAAATAAAGAAAGACCTCGGCTT 151

```

Db	47694	ATATAATTAATAATTTTATTATTAATAATAATTAATAATTAATAATAATAATAA	47753
QY	152	AACAGTCGAAAAACGAGAAATAATAAAAAGAGAGACTGCGATTTTAAATGGAATCGT	211
Db	47754	ATAAAATATATAAAACCAATTACTAAAAATAAATAATATATTTTAAATTAATAATAA	47813
QY	212	GAGAAAAGAAATTTTAAATTTTCATTTTCGAGGATTAATTTGTTGTAAGTGTAGTAA	271
Db	47814	AATTAATAATTAATAATAAATGCTTAATAAATGAAATAAAAAAT-----TATAATAATAA	47867
QY	272	AATCTAGATTAATAAATGACATCAAAATGTGTGAATTTGCATTTTGAATATGAAATACGTAG	331
Db	47868	TAGATAAATAATTAATAATCAAAAAATATACTAATAAATAATATAATCTAATAATAATA	47927
QY	332	TATATCAATATAGGGGTTTGTCTAATTTTATTTTTCGCAAGANTGAAATCTGAGTGAAG	391
Db	47928	ACATATAATTAATAATAATAATAATAATTAATAATGAAATAATACAAATAATAATAACT	47987
QY	392	AAATATCTTTCGAGAGCAAAAAACCCTTCGCGCTTTTTCGAAATGACTTTGGAAAAA	451
Db	47988	AGTCTAATAATAATAATTAATAATAATAATAATACTAATAATAATAATAATAATAA	48047
QY	452	TTCTATGTGACGCGTACGCAACTTTGAATTTTTCATATTCGAAATTTGAAAAATTAAG	511
Db	48048	ATATTTAAATTAATTAATAATTAATTAAGATAAATAATTAATCAATTAA---TAAAAATAA	48104
QY	512	GCAAAAGAACTCAATGAAAAAATATTTATTTTAAAAAAGAGATCGGATATGGAATT	571
Db	48105	AATTAATAATAATAATTAATGAAAAATAATTAATAATTAATAATAATAATAATAATAA	48164
QY	572	TAAAGAGAGAACTGCATTTGATGATGAAAAAAGATTTGGAAAAAATCTAGTCGAGAG	631
Db	48165	TTTAATTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	48224
QY	632	TGATTTTAAAGCAAAAAAATTTGGGAAGCTACACCGGCTGTGTTTGAATGACATGAA	691
Db	48225	TAAACACAGCTAATAATAATAATAATAATTAATTAATAATAATAATTAATAATAA	48284
QY	692	AGAAATGATGAAAAATATCAACTCGATGTTTAATGAGAAACGACAGTCAATATAGAA	751
Db	48285	AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	48344
QY	752	ATTATTTAAAAATTAAGAGTGAAGATAGTAATGATGATCTAGTACGGACTATTACA	811
Db	48345	TTTTTTTTTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	48404
QY	812	GTGTGGAAGTCGAATAGTTTGTATTAACAAGACAGCGCTTTCGTTTGTATTTGTGA	871
Db	48405	ATTAAGTAA--AATATATATATAGGAAAAAATTAACAAAAACAATTAATAATTAATAA	48462
QY	872	GAGATTTACAGCACTGAGAAAGAGAGCTGATTAATGCAAGAGAGTAAAAAGATTCGATAC	931
Db	48463	AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAC	48522
QY	932	AATGAAAGCAAAACCTAAGAGGCTTTTGAATTTGAGTTTGTGTTTGTATTAAGATTTTT	991
Db	48523	TACTCTGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA	48582
QY	992	GAGTGAATAATAGAAATCAATGGAATGATTTTTCACAACAAAAAAGACTCGCAAGTAA	1051
Db	48583	GAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATACTAT	48642
QY	1052	AAGAAAA 1058	
Db	48643	AATTAATA 48649	

VERSION	E23356.1	GI:13024379
KEYWORDS	JP 1999075859-A/2.	
SOURCE	unidentified.	
ORGANISM	unidentified	
REFERENCE	1 (bases 1 to 6644)	
AUTHORS	Hirofumi, H.	
TITLE	Virus vector system expressing apoptosis-related gene	
JOURNAL	Patent: JP 1999075859-A 2 3 -MAR-1999;	
COMMENT	R B R JENSEN KK OS Unidentified PN JP 1999075859-A/2 PD 23-MAR-1999 PF 08-SEP-1997 JP 1997259235 PR	
	HIROFUMI HAMADA PC C12N15/09, C12N5/10, C12N7/00, A61K35/76, A61K48/00, C12N5/10, PC C12R1:91), PC (C12N7/00, C12R1:92), C12N15/00, C12N5/00, (C12N5/00, C12R1:91) CC Strandedness: Double; CC Topology: Linear; FH Key Location/Qualifiers FT source 1..6644 FT /organism='Unidentified'.	
FEATURES	Location/Qualifiers 1..6644 /organism='unidentified'	
BASE COUNT	2166 a 1573 c 1424 g 1481 t	
ORIGIN		
Query Match	5.6% Score 93.6; DB 6; Length 6644;	
Best Local Similarity	45.9% Pred. NO. 8.5e-05;	
Matches 321; Conservative	0; Mismatches 379; Indels 0; Gaps 0;	
OY	7 ACACAGCGCAATTGAAAAAATACATATAAATTTTCGAGATTTAAAGCCGACTTAAAC	66
DB		3846
OY	67 AAATGAGTGAAGAAGAAAGAAAAATTAATCATATTTTTCAGTGTAGTAAAGAGAAG	126
DB		3906
OY	127 AAAAAATTAAGAAGACCTCGGCTTAACAGTCAAAAAACGAATAATAAAGAAAGA	186
DB		3966
OY	3907 AAAAAAATTAAGAAGACCTCGGCTTAACAGTCAAAAAACGAATAATAAAGAAAGA	186
DB		3966
OY	187 GACGTGATTTTAAATGGAATCGTGAAGAAAGAAAAATTTTAAATTTTCATTTTCGAGCG	246
DB		4026
OY	3967 AAAAAAATTAAGAAGACCTCGGCTTAACAGTCAAAAAACGAATAATAAAGAAAGA	186
DB		4026
OY	247 ATTAATTTGTTGTAAGTGAAGAAATCTGATTAATAAATCGATCAAAATGTGTG	306
DB		4086
OY	4027 AAAAAAATTAAGAAGACCTCGGCTTAACAGTCAAAAAACGAATAATAAAGAAAGA	186
DB		4086
OY	307 AATTGACATTTGAAATACGTATATCATATATGCGGTTGCTATTTTATTTTG	366
DB		4146
OY	4087 AAAAAAATTAAGAAGACCTCGGCTTAACAGTCAAAAAACGAATAATAAAGAAAGA	186
DB		4146
OY	367 CGAAGATTTGAATACTGAGTGAAGAAATATGTTGCGAGCAAAAAACCTTGCGCT	426
DB		4206
OY	4147 AAAAAAATTAAGAAGACCTCGGCTTAACAGTCAAAAAACGAATAATAAAGAAAGA	186
DB		4206
OY	427 TTTTTCAAATGACTTTGAAAAATTCATTTGAGCGGTAGCGAACTTTGAATTTT	486
DB		4266
OY	4207 AAAAAAATTAAGAAGACCTCGGCTTAACAGTCAAAAAACGAATAATAAAGAAAGA	186
DB		4266
OY	487 TACATTTGGAATTTGAAAAATTAAGCAAAAGAACTCAATGGAATAATATTATTATA	546
DB		4326
OY	4267 AAAAAAATTAAGAAGACCTCGGCTTAACAGTCAAAAAACGAATAATAAAGAAAGA	186
DB		4326
OY	547 AAAAAAGAGATCGGATTTGATTTTAAAGCAAAAAACGTGACTTGAATGAATAAAG	606
DB		606

Dd	4327	AAA	4306
Oy	607	ATTGGAAAATAATCTAGTGTGAGACTGAATTAAACCAAAAAATTTGGAACTCAACC	666
Dd	4387	AAAGAATTGTCGAGCTTGCC	4446
Oy	667	GTCGTGTTTTGAAATGCAGTGAAGAAATGATGAAAA	706
Dd	4447	GAGATTTTCAGACGACTAAGAACGTMAAATGCGAGAAAAA	4486
RESULT 12			
E23357		7372 bp	Linear
LOCUS	E23357	Virus vector system expressing apoptosis-related gene.	PAT 18-JUN-2001
DEFINITION	E23357		
ACCESSION	E23357.1	GI:13024380	
VERSION	JP 1999075859-A/3.		
KEYWORDS	Homo sapiens.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE			
AUTHORS	Mumukshu; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
JOURNAL	1 (bases 1 to 7372)		
COMMENT	Hirofumi, H.		
	Virus vector system expressing apoptosis-related gene		
	Patent: JP 1999075859-A 3 23-MAR-1999;		
	R B R JENSEN KK		
	OS Homo sapiens (human)		
	PN JP 1999075859-A/3		
	PD 23-MAR-1999		
	PF 08-SEP-1997 JP 1997259235		
	PR		
	PI HIROFUMI HAMADA		
	PC C12N15/09, C12N5/10, C12N7/00//A61K35/76, A61K48/00, (C12N5/10, PC		
	C12R1:91),		
	PC (C12N7/00, C12R1:92), C12N15/00, C12N5/00, (C12N5/00, C12R1:91) CC		
	Strandedness: Double;		
	CC Topology: Linear;		
	FH Key		
	FT source		
	Location/Qualifiers		
	1..7372		
	/organism='Homo sapiens'		
	/db_xref='taxon:9606'		
BASE COUNT	2353 a 1749 c 1649 g 1621 t		
ORIGIN			
	Query Match	5.6%; Score 93.6; DB 6; Length 7372;	
	Best Local Similarity	45.9%; Pred. No. 8.3e+05;	
	Matches 321; Conservative	0; Mismatches 379; Indels 0; Gaps 0;	
Oy	7	ACACAAGCCAATTGAAAAAATTAATATTTTCGCCGATTAAGCCGACTTAAAC	66
Dd	4515	AAA	4574
Oy	67	AAATGAGTAGAAGAAAGAAAAAATAATTAATCATATTTTGAGTTAGTAAGAAGAAG	126
Dd	4575	AAA	4634
Oy	127	AAAAAATTAAGAAGACCTCGGCTTAACAGTCGAAAAACGAATAATTAAGAAGAAGA	186
Dd	4635	AAA	4694
Oy	187	GACGTGTGTTTTTAATGGAATCGAGCGAAGAAAGAAATTTAAPTTTCATTTTCGAGCG	246
Dd	4695	AAA	4754
Oy	247	ATTAATTTTGTGTAGTGAATAATCTAGATTAATAAATGCAGATCAAAATGTGTG	306
Dd	4755	AAA	4814
Oy	307	AATTTGACATTATTGAAATACGATATATCAATTAATGCGGTTTGTCTATTTTATTTTG	366

Db		4815	AAA	4874
OY		367	CGAAGATTGAAATCTGAGTCAAGAAAATAGTTTGCGAGACGCAAAAAACCCTGCCGT	426
Db		4875	AAA	4934
OY		427	TTTTTCAAATGACTTGGAAAAATTCATGTGAGCGGTAGCGAACTTGAATTTT	486
Db		4935	AAA	4994
OY		487	TACATTGGAATTTGAAAAATAAGCAGAAAGAACCTCAATGCAGAAAAATATTATTATA	546
Db		4995	AAA	5054
OY		547	AAAAAGAGATCGGATATGGATTTTAAACGACAAACTGACATTGAATGAAAAAAGA	606
Db		5055	AAA	5114
OY		607	ATTTGAAAAAATCTATGCTGAGAGTCAITTAAGCAAAAAATTGGGAATCAACCCG	666
Db		5115	AAGATTGGT	5174
OY		667	GTCGTGTTTAGAATGACGTGAAGAAATGATGAAAAA	706
Db		5175	GAGATTTTCAGAGACTAAGGAGCTAAATATGAGAAAAAA	5214

[illegible]

RESULT 13	
LOCUS	E23355
DEFINITION	7797 bp DNA linear PAT 18-JUN-2001
ACCESSION	E23355
VERSION	E23355.1 GI:13024378
KEYWORDS	JP 1999075859-A/1.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 7797)
AUTHORS	Hirofumi H.
TITLE	Virus vector system expressing apoptosis-related gene
JOURNAL	Patent: JP 1999075859-A 1 23-MAR-1999;

Accession	Sequence	Position
Db	5420 AA	5479
Qy	547 AAAAAAGAGATCGGATATGGATTTTAAAGCAGAAAACTCACATTGATGAAAAAAG	606
Db	5480 AA	5539
Qy	607 ATTTCGAAATAATCTATGCTGAGAGATTAATTAAGCAAAAAATTGGGAATCAACCG	666
Db	5540 AAAGATTGGTCGACCTTGC	5599
Qy	667 GTCTGTTTGAAGATGACCATGAAACAAATGATGAAAA	706
Db	5600 GAGATTTTCAGGAGCTAAGGAGCTAAATATCGACAAATTA	5639

```

PF      08-SEP-1997 JP 1997259235
PR
PI      HIROFUMI HAMADA
PC      C12N15/09,(C12N5/10,C12N7/00//A61K35/76,A61K48/00,(C12N5/10,PC
        C12R1/91),
PC      (C12N7/00,C12R1/92),C12N15/00,C12N5/00,(C12N5/00,C12R1/91) CC
Strandedness: Double;
CC      Topology: Linear;
FH      key
FH      Location/Qualifiers
FT      source
        Location/Qualifiers
            1..7797
            /organism="Bovine poxvirus",
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT      2542 a 1760 c 1656 g 1839 t
ORIGIN

```

LOCUS	E23359	7996 bp	DNA	linear	PAT 18-JUN-2001
DEFINITION	Virus vector system expressing apoptosis-related gene.				
ACCESSION	E23359				
VERSION	E23359.1	GI:13024382			
KEYWORDS	JP 1999075859-A/5.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	I (bases 1 to 7996)				
TITLE	Hirofumi, H.				
JOURNAL	Virus vector system expressing apoptosis-related gene				
	Patent: JP 1999075859-A 5 23-MAR-1999;				
	R. B. R. JENSEN KK				
COMMENT	OS Homo sapiens (human)				
	PN JP 1999075859-A/5				
	PD 23-MAR-1999				

Query Match	5.6%	Score 93.6	DB 6	Length 7797
Best Local Similarity	45.9%	Pred. No. 8,2e-05		
Matches	321	Conservative	0	Mismatches 379; Indels 0; Gaps 0;
QY	7	ACACACAGCAATTGAAAAAATAGATAAATTTTCGACAGGTATTAAGCCGACTTAAAC	66	
Db	4940	AA	49999	
QY	67	AAATGACTGAGAGAAAGAAAAAATTAATACATATTTTGACTAGTAAGAGAGAAG	126	

PR	HIROFUMI HAMADA
P1	.C12N15/09,.C12N5/10,.C12N7/00/.A6IK35/76,A6IK48/00,(.C12N5/10, PC .C12R1:91)
PC	(.C12N7/00,.C12R1:92),.C12N15/00,.C12N5/00,(.C12N5/00,.C12R1:91) CC
Strandedness:	Double;
CC	Topology: linear;
FH	Key
FT	Source Location/Qualifiers 1..7996

FEATURES FT /Organism='Homo sapiens (human)'
 source 1. 7996 /Organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 2463 a 2015 c 1829 g 1689 t
 ORIGIN

Query Match 5.6%; Score 93.6; DB 6; Length 7996;
 Best Local Similarity 45.9%; Pred. No. 8.1e-05;
 Matches 321; Conservative 0; Mismatches 379; Indels 0; Gaps 0;

QY 7 ACACAGGCGCAATTTGAAATAGATAATTTCCAGAGTATTAAGCCGACTTAAAC 66
 DB 5139 AA 5198
 QY 67 AATGAGTGAAG 126
 DB 5199 AA 5258
 QY 127 AAAAAATTAAG 186
 DB 5259 AA 5318
 QY 187 GACTGTGATTTTATGAAATCGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
 DB 5319 AA 5378
 QY 247 ATTAAATTTGTTAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306
 DB 5379 AA 5438
 QY 307 AATTGACATTTATGAAATCTAGATATCATATATGAGGCTTTGCTATTTATTTTG 366
 DB 5439 AA 5498
 QY 367 CGAAGATTTGAAATCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426
 DB 5499 AA 5558
 QY 427 TTTTTCGAATGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
 DB 5559 AA 5618
 QY 487 TACATGGAATTTGAAATTAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
 DB 5619 AA 5678
 QY 547 AAAAAAGAGATCGATATGATTTTAAAGCAGAGAGAGAGAGAGAGAGAGAGAG 606
 DB 5679 AA 5738
 QY 607 ATTGAG 666
 DB 5739 AA 5798
 QY 667 GTGTGTTTATGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
 DB 5799 GAGATTTTCAG 5838

RESULT 15
 AC116966/c 268147 bp DNA linear HTG 04-APR-2002
 LOCUS
 DEFINITION Dictyostellium discoideum chromosome 2 map 2482102-2750247 strain
 AX4, *** SEQUENCING IN PROGRESS ***; in ordered pieces.
 ACCESSION AC116966
 VERSION AC116966.1 GI:19920065
 KEYWORDS HTG: HTGS PHASE2.
 SOURCE Dictyostellium discoideum.
 ORGANISM Dictyostellium discoideum
 Eukaryota; Eukaryota; Dictyostellida; Dictyostellium.
 REFERENCE 1 (bases 1 to 268147)

AUTHORS Gloeckner, G., Eichinger, L., Szafarski, K., Pacheco, J., Dear, P.,
 Lehmann, R., Baumgart, C., Parra, G., April, J.F., Guigo, R., Kumpf, K.,
 Tunggal, B., Cox, E., Quail, M.A., Platzer, M., Rosenthal, A. and
 Noegel, A.A.
 TITLE Sequence and Analysis of Chromosome 2 of Dictyostellium
 JOURNAL Unpublished
 REFERENCE The Dictyostellium Genome Sequencing Consortium
 2 (bases 1 to 268147)
 AUTHORS Baumgart, C.
 TITLE Direct Submission
 COMMENT Submitted (04-APR-2002) Genome Analysis, Institute of Molecular
 Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
 CDS predictions from GenBank may contain errors. Further information
 is available from IMB Jena, Department of Genome Analysis
 (http://genome.imb-jena.de/dictyostellium/)
 and the University Cologne, Institute for Biochemistry I
 (http://www.uni-koeln.de/dictyostellium/project.shtml)
 Funding
 Agency : Deutsche Forschungsgemeinschaft (DFG).
 * NOTE: This is a 'working draft' sequence.
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source 1. 268147
 /Organism='Dictyostellium discoideum'
 /strain='AX4'
 /db_xref='taxon:44689'
 /Chromosome='2'
 /map='2482102-2750247'
 BASE COUNT 101633 a 30111 c 30461 g 105740 t 202 others
 ORIGIN

Query Match 5.6%; Score 93.4; DB 2; Length 268147;
 Best Local Similarity 45.5%; Pred. No. 3.6e-05;
 Matches 409; Conservative 0; Mismatches 486; Indels 4; Gaps 2;

QY 6 TACACAGCGCAATTTGAAATAGATAATTTCCAGAGTATTAAGCCGACTTAAAC 65
 DB 197289 TAAAAAAAAAAAAAAAAAAAAAAAAACAGAGAGAGAGAGAGAGAGAGAGAG 197230
 QY 66 CAATGAGTGAAG 125
 DB 197229 AA 197170
 QY 126 GAAAAATTAAG 185
 DB 197169 AA 197110
 QY 186 AGACTGATTTTATGAAATCGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 245
 DB 197109 AAAAAAAAAACCTTTTGTGAGTCTTTATTAATTAATTAATTAATTTT 197050
 QY 246 GATTAAATTTGTTAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 305
 DB 197049 TTTTGAATTTATTTATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196990
 QY 306 GAATTTGACATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 365
 DB 196989 - - - ATTCATATGTTTATTAATTCACACCCCTCAATTTGTTAAATTAAGAGAGAGAG 196933
 QY 366 GCGAGATTTGAAATCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 425
 DB 196932 ATTAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 196873
 QY 426 TTTTTCGAATGACTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
 DB 196872 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 196813
 QY 486 TTACATTTGAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 545
 DB 196812 TATCACTATTAAGCTCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 196753

